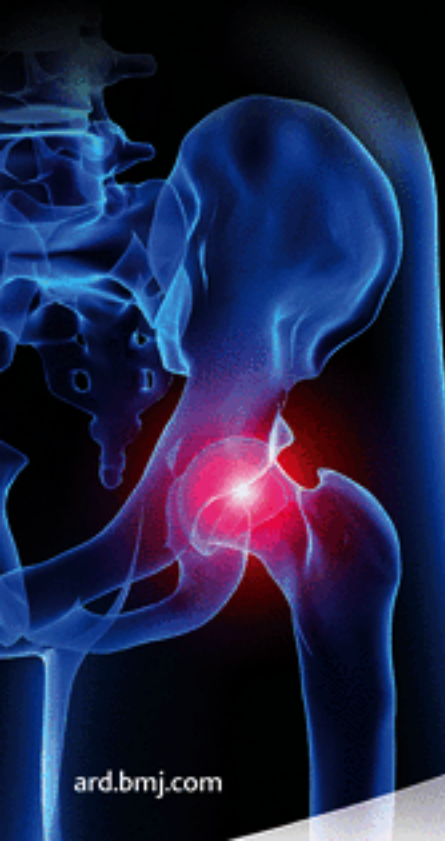
March 2020 Volume 79 Issue 3

Annals of the Rheumatic Diseases

The EULAR Journal



The home for innovative rheumatology research

ard.bmj.com

eular BMJ

ARD
The EULAR Journal

Guidelines for Authors and Reviewers

Full instructions are available online at http://ard.bmj.com/pages/authors. Articles must be submitted electronically at http://mc.manuscriptcentral.com/ard. Authors retain copyright but are required to grant ARD an exclusive licence to publish. (http://ard.bmj.com/pages/authors/).

Annals of the Rheumatic Diseases publishes original work on all aspects of rheumatology and disorders of connective tissue. Laboratory and clinical studies are equally welcome

Subscription Information

ARD is published monthly; subscribers receive all supplements
ISSN 0003-4967 (print); 1468-2060 (online)

Institutional Rates 2020

Print
£1088

Online
Site licences are priced on FTE basis and allow access by the whole institution. Details available online at http://journals.bmj.com/content/subscribers or contact the Subscription Manager in the UK (see above right)

Personal Rates 2020

Print (includes online access at no additional cost)
£450

Online only
£192

EULAR congress delegates
Delegates receive a Continuous Professional Development package that includes a 12 month complimentary subscription to *ARD* in print and/or online

Personal print or online only and institutional print subscriptions may be purchased online at http://journals.bmj.com/content/subscribers (payment by Visa/Mastercard only)

Residents of some EC countries must pay VAT; for details, call us or visit http://journals.bmj.com/content/subscribers

For more information on subscription rates or to subscribe online please visit ard/bmj.com/pages/contact-us/

Contact Details

Editorial Office
Annals of the Rheumatic Diseases
BMJ Journals, BMA House, Tavistock Square
London WCIH 9JR,UK
E: **ard@bmj.com**

Production Editor
Teresa Jobson
E: **production.ard@bmj.com**

EULAR
EULAR Executive Secretariat
Seestrasse 240, 8802 Kilchberg, Switzerland
E: **eular@eular.org**
www.eular.org

Customer support
For general queries and support with existing and new subscriptions:
W: **support.bmj.com**
T: **+44 (0)20 7111 1105**
E: **support@bmj.com**

Self-archiving and permissions
W: **bmj.com/company/products-services/rights-and-licensing/**
E: **bmj.permissions@bmj.com**

Advertising
W: **bmj.com/company/for-advertisers-and-sponsor/**

Display Advertising ROW
Sophie Fitzsimmons
T: **+44 (0)20 3655 5612**
E: **sfitzsimmons@bmj.com**

Online Advertising ROW
Marc Clifford
T: **+44 (0) 20 3655 5610**
E: **mclifford@bmj.com**

Display & Online Advertising Americas
American Medical Communications (AMC)
T: **+1 973 214 4374**
E: **rgordon@americanmedicalcomm.com**

Reprints

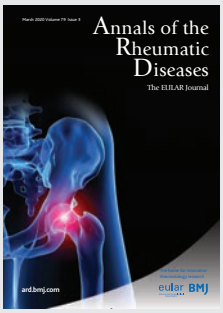
Author Reprints
BMJ Reprints Team
E: **admin.reprints@bmj.com**

Commercial Reprints ROW
Nadia Gurney-Randall
M: **+44 07866 262344**
E: **ngurneyrandall@bmj.com**

Commercial Reprints Americas
Ray Thibodeau
T: **+1 267 895 1758**
M: **+1 215 933 8484**
E: **ray.thibodeau@contentednet.com**

For all other journal contacts
ard.bmj.com/contact-us

ARD
The Eular Journal
Impact Factor: 14.299



Editorial office
Annals of the Rheumatic Diseases
BMJ Publishing Group Ltd
BMA House
Tavistock Square
London WCIH 9JR,UK
T: **+44 (0)20 3655 5889**
E: **ard@bmj.com**
Twitter: **@ARD_BMJ**
ISSN: 0003-4967 (print)
ISSN: 1468-2060 (online)

ARD is published by BMJ Publishing Group Ltd typeset by Exeter Premedia Services Private Ltd, Chennai, India and printed in the UK on acid-free paper.

Annals of the Rheumatic Diseases (ISSN No: 0003-4967) is distributed in the USA by Air Business Ltd. Periodicals postage paid at Jamaica NY 11431. POSTMASTER: send address changes to *Annals of the Rheumatic Diseases*, Air Business Ltd, c/o WN Shipping USA, 156-15, 146th Avenue, 2nd Floor, Jamaica, NY 11434, USA.

MORE CONTENTS ►

Myositis

Imaging

Osteoarthritis

Crystal arthropathies

Letters

Electronic pages

Chronic inflammatory arthritis following checkpoint inhibitor therapy for cancer: game changing implications

Leonard Calabrese ,[1] Xavier Mariette[2,3]

The use of immune checkpoint inhibitor (ICI) therapy for cancer is now a pillar of oncological therapeutics and growing, with an estimated 43.5% of all tumours falling within current labelling indications for use.[1] Eventually as the accessibility to ICI therapy increases, these data have staggering implications, given that an estimated number of new cancers in Europe and the USA exceeds 5 million individuals yearly.[2] As a byproduct of this tidal wave of newly exposed patients to various forms of immunotherapy with estimates that 10%–20% or more who may develop serious immune related adverse events (irAEs),[3] it is inevitable that the evaluation and care of such patients will pose a challenge to existing healthcare systems and likely create a space for a new specialty to manage such. From a rheumatological perspective, let us now consider that an estimated 3%–7% of ICI exposed patients may develop inflammatory arthritis (IA),[4 5] making it seem inevitable that ICI associated IA will become ever more commonplace, giving us pause to ask ourselves what our current understanding of this disorder is and how prepared we are to manage it.

These irAEs are heterogeneous and appear to differ in their presentations, similarity to existing constructs of autoimmune diseases and their natural history. To help focus the discussion regarding these complications and based on the available data on IrAEs, we propose a classification of them into three main categories (box 1). Most irAEs are self-limiting in nature and while they may have lasting clinical effects such as ongoing requirement for hormone replacement therapy in some endocrinopathies, the inflammatory phase of these illnesses is largely self-limiting with few exceptions, with less than 10% requiring additional therapy after suppression with glucocorticoids.[6] A second category is the development of a classical autoimmune disease in subjects who were predisposed. Indeed, many of these patients have been documented to have pre-existing specific autoantibodies (anti-cyclic citrullinated peptide (CCP) and they develop rheumatoid arthritis (RA), anti-SSA and they develop Sjögren's syndrome) and ICI therapy appears to act as a trigger the underlying autoimmune disease.[7]

A third category could be that described in the current report by Braaten and colleagues,[8] as they report persistent and ongoing non-specific IA. Until now, there has been a scant literature on this type of prolonged irAE, which makes this study of interest and importance. In their study published in *The Annals*, they interrogated a prospective observational data base from their centre of all patients referred with IA associated with ICIs and focused on those who had persistent arthritis for up to 24 months after ICIs had been stopped for treatment completion, disease progression or toxicity. With a mean follow of 9 months, they observed a remarkable 53% had active IA at last follow-up with one-quarter with active disease at 24 months. The vast majority (80%) were treated with some dose of glucocorticoids and 24 patients required disease modifying anti-rheumtic drugs (DMARDs) including 11 with biologic disease modifying anti-rheumatic drugs (bDAMRDs). Interesting trends were also noted in that these patients appeared more likely to develop chronic IA if they had longer exposure to ICIs and had a history of other forms or irAEs as well. Based on these data and hints from previous smaller reported experiences, it appears that IA may be the first irAE associated with a high likelihood of developing into a chronic autoimmune and or autoinflammatory complication of ICI therapy. Clearly larger numbers of patients studied over longer time period will be required to quantify this assertion.

Based on our current understanding of IA as an irAE and its propensity for chronicity we should question what this disease truly represents, specifically asking whether it is a traditional form of IA such as RA, spodyloarthritis (SpA) or other condition or alternatively does it represent a new nosological entity in itself. As in previous reports, their patients were vastly seronegative with a tendency for more pauciarticular disease over polyarticular ones.[3] In terms of etiopathogenesis,

Box 1 Proposed classification of immune related adverse events (irAEs)

Type 1
irAEs that are self-limiting in their inflammatory phase either though use of short-term immunosuppression or discontinuation of immune checkpoint inhibitors. This is the most common pattern reported. Type 1 reactions are generally non-specific in nature and not consistent with traditional classifications of autoimmune diseases.

Type 2
irAEs that appear indistinguishable for idiopathic forms of autoimmune diseases and often identified by the presence of signature autoantibodies such as antibodies to citrullinated proteins in rheumatoid arthritis, anti-acetylcholine esterase antibodies in myasthenia gravis, anti-islet cell antibodies in type 1 diabetes.[3] Type 2 irAEs are rare and represent only a small proportion of all irAEs. These tend to be chronic but the natural history is still poorly characterised.

Type 3
irAEs that are chronic in their inflammatory phase and based on the current literature and the report by Brateen *et al*,[8] inflammatory arthritis (IA) appears to be the most common irAE to assume this clinical course. Rare reports of chronic and or relapsing colitis and pneumonitis and dermopathy[21] have been reported but in general aside from IA descriptive reports are rare.

[1]Rheumatology/Immunology, Cleveland Clinic, Cleveland, Ohio, USA
[2]Rheumatology, Assistance Publique-Hôpitaux de Paris (AP-HP), Hôpitaux universitaires Paris-Sud – Hôpital Bicêtre, Le Kremlin Bicêtre, France
[3]AP-HP Université Paris Saclay, Center for Immunology of Viral Infections and Auto-immune Diseases (IMVA), Institut pour la Santé et la Recherche Médicale (INSERM) UMR 1184, Université Paris-Saclay, Le Kremlin Bicêtre, France

Correspondence to Dr Leonard Calabrese, Rheumatology/Immunology, Cleveland Clinic, Cleveland, OH 44106, USA; calabrl@ccf.org

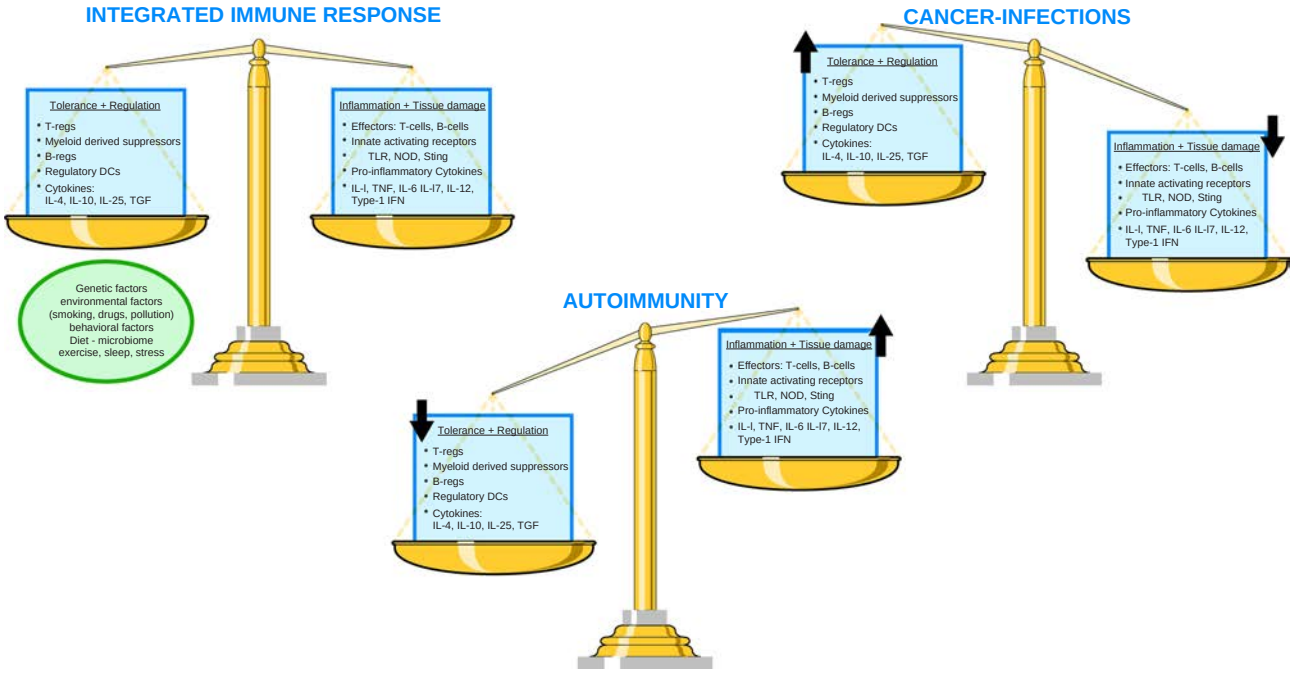


Figure 1 Immune balance. The integrated immune response in its homoeostatic state is a balance of complex tolerogenic and inflammatory forces, each contributing to optimal surveillance and patrolling functions designed to detect and dispatch danger while preserving the integrity of the host. The end product of the response is influenced by a variety of factors including the hosts genetic background and external variables and events that potentially can alter the immune system. In cancer and chronic infections there is an imbalance of depressed local-regional or systemic effector functions allowing continued growth of the tumour or persistence of infection. In autoimmunity, inflammatory effector function is overactive relative to tolerance and regulation. Both immunotherapy with ICIs and immunosuppressive therapies can shift this balance, with immune checkpoint inhibitors increasing local/systemic inflammatory reactions and immunosuppressive therapies such as glucocorticoids and cDMARDS and bDMARDS and other therapies suppressing inflammation with the theoretical possibility of compromising antitumoural effects. DC, dendritic cells; IL, interleukin; NOD, nucleotide-binding oligomerisation domain-like receptors; STING, stimulator of IFN genes; TLR, toll like receptors.

the picture is far from clear but despite the preponderance of the absence of antibodies to citrullinated proteins and rheumatoid factor, a recent investigation[9] has revealed that shared epitope alleles were more common in patients with ICI associated IA suggesting some commonality with RA. Clearly more investigation at the basic and translational level is needed to gain a better understanding of the condition.

The management of irAEs remains a challenge for oncologists, rheumatologists and other specialists. There is now a consensus that when clinically confronting irAEs, except in severe cases of myositis, myocarditis, pneumonitis or inflammatory bowel disease, the primary objective is to facilitate optimal treatment of the underlying cancer with ICIs or other immune based therapies whenever possible.[10] The arbiter of this strategy is to balance the anti-inflammatory/immunosuppressive effects of the chosen therapy for the irAE while preserving the enhanced antitumoural effects of the immunotherapy[11] (figure 1). Steroids are first line therapy used with the objective of employing the smallest effective dose. There are ongoing debates about a possible deleterious effect of steroids on cancer response to ICIs. There are some evidence that when glucocorticoids are given at the time ICI therapy is commenced there appears to be an attenuation of their antitumoural effects.[12] In addition, limited data also suggest, that at least in one complication (ie, hypohysitis), high dose glucocorticoids may compromise antitumoural immunity.[13] At present however, there are no convincing data that that low doses (ie, less than 10 mg prednisone daily) are detrimental in this setting. It also should be noted that there are mixed data on the influence of the presence of any irAE on antitumoural immunity with some suggesting enhanced effects while others no effect.[14] It is interesting to note that in the current report by Braaten and colleagues,[8] patients with persistent IA tended to have a better cancer prognosis than patients without persistent IA, even if most of them were treated with steroids or DMARDS. Similar beneficial effects on cancer outcome of musculoskeletal irAEs has been previously described.[15 16] Further studies are necessary to determine if this new form of prolonged non-specific IA is specially associated with a better cancer prognosis.

For achieving the objective of tapering and stopping steroids, the use of DMARDs is recommended.[17] In the current report by Braaten and colleagues,[8] DMARDs were used in 40% of ICI-induced IA, including bDMARDs in almost half of them. In their study, use of steroids, classical synthetic disease modifying anti-rheumatic drugs (csDMARD) or bDMARD did not impact cancer outcomes. Obviously, the discussion of using DMARDs or not is particularly important in the context of this type of irAE, that is persistent non-specific IA (type 3). Overall the approach must be based on the balance between benefits and risks and may evolve depending on future studies.

The use of bDMARD to treat irAEs may raise a series of general concerns to the rheumatologist. These targeted therapies have been licensed for more than 20 years and there has been a justified anxiety on behalf of patients and providers about a possible increased risk of cancer with the class. Moreover, abatacept, a highly effective targeted therapy is the converse of ipilimumab, a very efficient ICI. However, in spite of an alert on a possible increased risk of cancer, and especially of lymphoma, with monoclonal antitumour necrosis factor (TNF) antibodies in 2006, there is no convincing evidence

of increased risk of cancer related deaths with any bDMARD, even when used in patients with pre-existing malignancies.[18]

An even more provocative concept to consider is the systematic use of some bDMARDs like TNF inhibitors, tocilizumab or other targeted therapies in association with ICIs for preventing severe irAEs like severe colitis. In a preclinical study in mice, this approach was successful for preventing colitis, but also was beneficial in terms of anti-ancer effect.[19] The same observation was observed in a mouse model of melanoma where the addition of anti-TNF to anti-PD1 improved cancer control and survival.[20] In this study, anti-TNF attenuated overexpression of TIMP3, another checkpoint induced by anti-PD1. The beneficial effect of this association might be linked to a deleterious effect of inflammation for the action of ICIfigure 2). It is well known that a milieu of chronic inflammation may favours T cells exhaustion[21] that may limit the effectiveness of ICI. Clinical studies are ongoing evaluating combination of ICIs with systemic or intratumoural use of TNF inhibitors or tocilizumab (wwwclinincaltrials.gov NCT03293784 and NCT03588936).

In conclusion, it appears that some irAEs, especially IA, may evolve into chronic and possibly permanent inflammatory diseases that will require ongoing and perhaps lifelong immunosuppressive therapy. Rheumatologists are already being called on to participate in the management of these complex and challenging patients and a new field of irAE medicine is evolving. In this context, it is possible that, paradoxically, combining TNF or interleukin 6 inhibitors to ICI could both avoid irAEs and increase the efficacy of ICI in some specific situations. Studies to clarify what will be optimal therapy to control ongoing inflammatory diseases while preserving antitumoural immune responses are urgently needed.

Handling editor Josef S Smolen

Contributors Both LC and XM participated in all aspects of writing and reviewing the work.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Commissioned; internally peer reviewed.

To cite Calabrese L, Mariette X. *Ann Rheum Dis* 2020;**79**:309–311.

Received 21 October 2019
Revised 14 December 2019
Accepted 16 December 2019
Published Online First 3 January 2020

Linked

► http://dx.doi.org/10.1136/annrheumdis-2019-216109

Ann Rheum Dis 2020;**79**:309–311.
doi:10.1136/annrheumdis-2019-216510

ORCID iD
Leonard Calabrese http://orcid.org/0000-0002-1789-4923

REFERENCES

1 Haslam A, Prasad V. Estimation of the percentage of US patients with cancer who are eligible for and respond to checkpoint inhibitor immunotherapy drugs. *JAMA Netw Open* 2019;2:e192535.
2 World Health Organization Europe. Data and statistics, 2019. Available: http://www.euro.who.int/en/health-topics/noncommunicable-diseases/cancer/data-and-statistics
3 Calabrese LH, Calabrese C, Cappelli LC. Rheumatic immune-related adverse events from cancer immunotherapy. *Nat Rev Rheumatol* 2018;14:569–79.
4 Cappelli LC, Gutierrez AK, Baer AN, *et al*. Inflammatory arthritis and sicca syndrome induced by nivolumab and ipilimumab. *Ann Rheum Dis* 2017;76:43–50.
5 Calabrese C, Kirchner E, Kontzias A, *et al*. Rheumatic immune-related adverse events of checkpoint therapy for cancer: case series of a new nosological entity. *RMD Open* 2017;3:e000412.
6 Shoushtari AN, Friedman CF, Navid-Azarbaijani P, *et al*. Measuring toxic effects and time to treatment failure for nivolumab plus ipilimumab in melanoma. *JAMA Oncol* 2018;4:98–101.
7 Belkhir R, Burel SL, Dunogeant L, *et al*. Rheumatoid arthritis and polymyalgia rheumatica occurring after immune checkpoint inhibitor treatment. *Ann Rheum Dis* 2017;76:1747–50.
8 Braaten TJ, Brahmer JR, Forde PM, *et al*. Immune checkpoint inhibitor-induced inflammatory arthritis persists after immunotherapy cessation. *Ann Rheum Dis* 2019. doi:10.1136/annrheumdis-2019-216109. [Epub ahead of print: 20 Sep 2019].
9 Cappelli LC, Dorak MT, Bettinotti MP, *et al*. Association of HLA-DRB1 shared epitope alleles and immune checkpoint inhibitor-induced inflammatory arthritis. *Rheumatology* 2019;58:476–80.
10 Calabrese L, Mariette X. The evolving role of the rheumatologist in the management of immune-related adverse events (irAEs) caused by cancer immunotherapy. *Ann Rheum Dis* 2018;77:162–4.
11 Bucktrout SL, Bluestone JA, Ramsdell F. Recent advances in immunotherapies: from infection and autoimmunity, to cancer, and back again. *Genome Med* 2018;10:79.
12 Arbour KC, Mezquita L, Long N, *et al*. Impact of baseline steroids on efficacy of programmed cell death-1 and programmed Death-Ligand 1 blockade in patients with non-small-cell lung cancer. *J Clin Oncol* 2018;36:2872–8.
13 Faje AT, Lawrence D, Flaherty K, *et al*. High-dose glucocorticoids for the treatment of ipilimumab-induced hypophysitis is associated with reduced survival in patients with melanoma. *Cancer* 2018;124:3706–14.
14 Postow MA, Sidlow R, Hellmann MD. Immune-Related adverse events associated with immune checkpoint blockade. *N Engl J Med* 2018;378:158–68.
15 Buder-Bakhaya K, Benesova K, Schulz C, *et al*. Characterization of arthralgia induced by PD-1 antibody treatment in patients with metastasized cutaneous malignancies. *Cancer Immunol Immunother* 2018;67:175–82.
16 Mitchell EL, Lau PKH, Khoo C, *et al*. Rheumatic immune-related adverse events secondary to anti-programmed death-1 antibodies and preliminary analysis on the impact of corticosteroids on anti-tumour response: a case series. *Eur J Cancer* 2018;105:88–102.
17 Puzanov I, Diab A, Abdallah K, *et al*. Managing toxicities associated with immune checkpoint inhibitors: consensus recommendations from the Society for immunotherapy of cancer (SITC) toxicity management Working group. *J Immunother Cancer* 2017;5.
18 Regierer AC, Strangfeld A. Rheumatoid arthritis treatment in patients with a history of cancer. *Curr Opin Rheumatol* 2018;30:288–94.
19 Perez-Ruiz E, Minute L, Otano I, *et al*. Prophylactic TNF blockade uncouples efficacy and toxicity in dual CTLA-4 and PD-1 immunotherapy. *Nature* 2019;569:428–32.
20 Bertrand F, Montfort A, Marcheteau E, *et al*. Tnfα blockade overcomes resistance to anti-PD-1 in experimental melanoma. *Nat Commun* 2017;8:2256.
21 Wherry EJ. T cell exhaustion. *Nat Immunol* 2011;12:492–9.

The earlier, the better or the worse? Towards accurate management of patients with arthralgia at risk for RA

Annette van der Helm-van Mil [1,2] Robert B M Landewé[3,4]

Handling editor Josef S Smolen

[1]Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
[2]Rheumatology, Erasmus Medical Center, Rotterdam, The Netherlands
[3]Amsterdam Rheumatology Center, AMC, Amsterdam, The Netherlands
[4]Rheumatology, Zuyderland MC, Heerlen, The Netherlands

Correspondence to
Professor Annette van der Helm-van Mil, Rheumatology, Leiden University Medical Center, Leiden, Netherlands; A.H.M.van_der_Helm@lumc.nl

Received 26 November 2019
Accepted 2 January 2020
Published Online First 8 January 2020

To cite: van der Helm-van Mil A, Landewé RBM. *Ann Rheum Dis* 2020;**79**:312–315.

ABSTRACT

The favourable long-term results of early treatment in patients with classified rheumatoid arthritis have resulted in an increasing interest in the diseases phases preceding clinical arthritis. The hypothesis to test is that an intervention in these early phases may better prevent or reduce disease persistence than an intervention when arthritis has become clinically manifest. While several placebo-controlled trials are still ongoing, to date there is no firm evidence that this hypothesis truly holds. Therefore, it is important to reflect on the current status of arthralgia preceding clinical arthritis. Inherent to every new field of research, attitudes are conflicting, with opinions propagating innovation (based on the fear of undertreatment) on the one hand, and critical sounds pleading for more restraint (fear of overtreatment) on the other hand. In this Viewpoint, we will examine these divergent opinions, relate them to a preferred ultimate scenario and provide considerations for future studies and daily practice.

INTRODUCTION

Early treatment start has become the cornerstone of the management of early arthritis and rheumatoid arthritis (RA).[1 2] This concept has resulted in an increasing interest in the period before swollen joints actually appear, with the underlying assumption that disease processes that are still developing are more susceptible to permanent modification. Treatment success in the prearthritis phase will for instance be reflected by a lower incidence of persistent clinical arthritis, as disease persistence can hardly be affected anymore when treatment is initiated only when arthritis has appeared. Whether disease modifying anti-rheumatic drug (DMARD) treatment started in the prearthritis phase is truly more effective will become clear from the currently ongoing placebo-controlled proof-of-concept trials. Results will become available in the next 2 years.[3–6]

Therefore, it is important now to reflect on the status of the field of arthralgia preceding clinical arthritis. How should we deal with the temptations of scientific progress? Should we already try to positively influence the lives of current individuals presenting with arthralgia, but still without arthritis? How will such a practice influence the chance that we will ever get real evidence-based treatments in this field that are beyond beliefs? In this Viewpoint, we will investigate these questions. We will search for a trade-off between opinions propagating innovation on the one hand and critical sounds pleading for restraint on the other hand, conflicting opinions that are inherent to new fields of research.

THE DILEMMA

In the absence of evidence for the value of drug interventions in the phase preceding clinical arthritis, several scenarios may apply to describe the current thinking.

The first is the scenario of hope. To cite the American philosopher J. Dewey ‘every great advance in science has issued from a new audacity of imagination’. Treatment of RA has considerably improved and the hope is to further improve the lives of patients with RA, curing the disease or even prevent it. This hope is fuelled by promising results of DMARD-intervention in early RA. The contribution of timing of treatment to treatment success has driven researchers to study biological mechanisms that precede clinical arthritis. In clinical practice, patients with arthralgia suspicious of progression to RA, but yet without clinically apparent arthritis, are increasingly recognised. Until recently, these patients were sent home with the advice to come back in case of clinical arthritis. To date, some clinicians feel the need to start DMARD-therapy prior to the development of clinical arthritis, especially if laboratory or imaging findings suggest that RA is looming. They hope to add value to the lives of these patients and argue that—when ongoing trials will teach us that treatment in the prearthritis phase is effective—withholding DMARD-treatment would in hindsight mean undertreatment, a situation that they would rather avoid.

The second scenario is that of criticism or perhaps pessimism. The underlying sentiment is that, now disease activity can be so well suppressed in most RA patients, too early treatment may do more harm than good. In other words, very early treatment start could result in ‘overdiagnosis’ and overtreatment. In this view, the valuable progress that has been achieved has created a new problem. The British psychologist Havelock Ellis described it as follows: ‘What we call progress is the exchange of one nuisance for another nuisance’.

As so often, a third scenario that appropriately balances the risks of undertreatment and overtreatment is likely the ideal scenario.

LOOKING BACK

Balancing the risks of undertreatment and overtreatment in RA is not new. Fifteen years ago, similar discussions pertained to patients with undifferentiated arthritis (UA), that is, patients with clinically evident arthritis that do not fulfil classification criteria for RA or other inflammatory arthritides. Initially, patients diagnosed with UA were not treated with DMARDs. Only after ample validation of models predicting the risks in individual patients with acceptable accuracy,[7–12] the

notion occurred that UA should better be treated. Finally, in order to be able to also classify patients earlier in time, novel classification criteria for RA were developed.[13] Intriguingly, all placebo-controlled randomised clinical trials in patients with UA were negative for their primary endpoint, the fulfilment of classification criteria for RA.[14–18] These negative results may be explained by methodological limitations, such as small sample sizes and absence of risk stratification at inclusion. They may also indicate that—if prevention of RA is the ultimate goal—the phase of UA is too late to start DMARD-treatment. Posthoc analysis of a trial with methotrexate in the subgroup of patients with a high risk to progress to RA showed statistically significant and clinically relevant effects.[19] A meta-analysis of all trials performed in patients with UA provided similar results that were statistically significant.[20] To what extent overtreatment or undertreatment exists in UA will forever remain unclear, because DMARD treatment of patients with UA is common practice now. In the absence of solid scientific evidence from clinical trials, treatment decisions are guided by clinical expertise and personal experiences.

WHAT CAN WE LEARN FROM THE RECENT PAST?

In light of the preferred third scenario, important learning points from the past 15 years include the need to conduct well-designed placebo-controlled clinical trials, the need to base our future actions on the results of these trials and to refrain from implementing anticipated results in daily practice that are not (yet) existent.

Adequate trial design means requirements for statistical power, eligible patients, preferred outcome(s) and follow-up duration. Sufficient statistical power seems a trivial requirement but is tricky, because statistical power depends on the difficult to estimate proportion of patients who will get the ultimate outcome (RA). The failure of previous clinical trials with either UA or arthralgia to meet their primary endpoint has been attributed to insufficient sample sizes.

Choosing the best primary outcome is also not straightforward. The primary outcome used so far was fulfilment of RA according to classification criteria. The question is whether this outcome best reflects added value to patients. Nowadays the involvement of patient partners in research has paid off and patients have indicated that current disease burden is mostly caused by pain, fatigue and functional impairments.[21] From this perspective, added value may better be expressed as the possibility to acquire symptom resolution and maintain a normal daily living (including work). There is also increasing pressure from society to spend healthcare resources more parsimoniously, especially in light of the risen drug expenses for RA. Furthermore, from a methodological perspective, the achievement of an outcome at a single point in time is not reflective of the subsequent disease course. More specifically, the occurrence of clinically apparent arthritis (or RA) at a single time point does not say anything about whether the disease will be self-limiting, whether remission will be rapidly achievable with first-line therapy, whether a DMARD-free status can be achieved over time or whether the disease will be persistently active or poorly responsive. Sustained DMARD-free remission may be a better outcome, as it includes a form of persistence in its definition.[22] Altogether, a long-term follow-up of patients included in pre-RA trials is required to evaluate if results are sustainable and valuable, which is challenging as trials generally tend to shorten the follow-up duration. Considerations are summarised in box 1.

Box 1 Considerations on treatment of patients with arthralgia suspicious for progression to RA

- There is no evidence that starting DMARD treatment in this disease phase is effective.
- Several proof-of-concept trials are currently ongoing.
- Subsequent trials require long-term follow-up to determine if outcomes (absence of clinical arthritis, absence of persistent arthritis, achieving DMARD-free status) are sustainable. These trials should include outcomes that reflect real value to patients, such as patient-reported symptoms, functional ability and workability.
- The EULAR definition of arthralgia suspicious for progression to RA confers a high sensitivity for RA development but only a moderate specificity. Adding information from other biomarkers is needed to further increase specificity.
- Currently, there is no validated risk stratification method to reliably estimate the risk of progressing to RA. Analyses on a combination of markers in relation to the natural disease course in all relevant longitudinal data-sets are needed to achieve this.
- In daily clinical practice, rheumatologists may wish to balance the risks of overtreatment and undertreatment in patients with arthralgia suspicious for progression to RA. However, absence of evidence on risk estimations and on efficacy currently favours a decision not to treat arthralgia with DMARDs in the absence of clinical arthritis. Furthermore, evaluation of the natural course will shed light on risks in the nearby future.

RA, rheumatoid arthritis.

IDENTIFYING PERSONS WITH ARTHRALGIA AT RISK FOR RA

Defining the population of patients with arthralgia but without clinical arthritis, and who are supposedly at risk for RA, is another crucial element. The risk influences the required sample size: at a similar power, a larger sample is required if the risk for RA is low or moderate, as compared with a scenario in which the risk is high. Moreover, overtreatment becomes more of an issue if the actual proportion that will develop RA is smaller. Apart from accurate, the risk estimation should be robust and validated in data from different centres and countries. Since identifying RA purely based on underlying biological markers is still impossible, a proper diagnosis must rely on a combination of features and pattern recognition. A combination of clinical symptoms and signs suggestive of future RA has been developed, resulting in the EULAR definition of arthralgia suspicious for progression to RA.[23] The clinical definition has shown to be highly sensitive when tested against the external standard ‘expert diagnosis of RA’ in cross-sectional studies and actual RA development in longitudinal studies.[23 24] In order to obtain a sufficiently high specificity, the clinical definition should be combined with the results of biomarkers. Accumulating evidence suggest that autoantibodies and imaging-detected subclinical inflammation are the most promising biomarkers. A non-systematic look in the literature yields several different prediction models that have been construed.[25–29] Unfortunately, none of these has reported cross-validation in independent centres, which leaves researchers with residual uncertainty. Collaboration based on data sharing across centres and obtaining consensus on the preferred methodology is needed to optimally define a population at high risk to be included in future trials.

ASSESSING ACCURACY OF IDENTIFYING RISK FOR RA

Risk estimators for groups of patients in trials differ principally from those for individuals in clinical practice. In individual patients, positive and negative predictive values, expressing the likelihoods of contracting a disease or remain free of that, are the pivotal estimates. Ideally, these should approach 100%. This is an almost impossible scenario in rheumatology. Diagnostic criteria that are intended to be used in individual patients are therefore no longer pursued by the professional organisations. Initially, prognostic research has focused on the ability to identify patients at risk for RA: the sensitivity. However, as a high sensitivity harbours the risks of 'overclassification' and overtreatment, attention has shifted from high sensitivity to high specificity (or properly recognising the persons that will *not* develop RA). The best risk classification model therefore has a high accuracy based on high sensitivity *and* high specificity. Designing and validating such a model is a *'herculean task'* since sensitivity and specificity tend to operate in opposite directions. A factor that further complicates the matter is that even in the presence of a high specificity absolute likelihoods can still be low (Bayesian rule). A good example is the ACPA-test, with a documented specificity of 98%, that in populations with a low prior risk of RA, such as the general population, yields an individual likelihood of RA development of only 5%, corresponding to a likelihood of 95% of not getting RA.[30 31] In more selected populations with higher prior risk, higher positive predictive values (PPVs) can be found.[28 32] 'The pre-RA period' is a continuum that extends from health to the time immediately before the development of clinical arthritis and diagnosis or classification of RA. The risk of persistent disease varies by the place in this spectrum; risk stratification algorithms should therefore be developed for subpopulations separately.

CONSIDERATIONS OTHER THAN ACCURACY

It is arguable, though, whether trials on DMARD-treatment in at risk populations are only justified in the context of optimally accurate prediction models. Whether overtreatment or undertreatment, due to suboptimal accuracy, will be considered socially acceptable depends on many factors such as the likelihood of harm (toxicity of treatment, psychological harm caused by uncertainty about getting ill), treatment expenses, and consequences of missing a diagnosis. Satisfactory answers can only be provided by international consensus about preferable risk stratification models, validation of such models in international databases with data about the natural course and all levels of variability. An estimation of the added value for individuals should be part of discussion. These discussions that involve all stakeholders may ultimately lead to consensus on what is the best trade-off between 'ideal' and 'feasible'. Importantly, persons at risk should be included in these discussions, as their beliefs and preferences will predict treatment uptake.[33] Optimal participation in this process requires that information is lucid, fair and comprehensible to lay-people.[34]

So far, we have focused on pharmacological interventions in selected populations, but we appreciate the relevance of generic lifestyle interventions such as smoking cessation. Such interventions have a lower risk of harm than DMARDs and are also associated with other positive public health effects.

WHAT DOES THE CURRENT SITUATION IMPLY FOR PATIENTS WITH ARTHRALGIA SUSPICIOUS FOR PROGRESSION TO RA IN DAILY PRACTICE?

Since there is no broadly accepted method to identify patients at risk for RA with sufficient precision, scenario three in which both undertreatment and overtreatment are minimised does not yet exist. As discussed in the previous paragraphs, long-term observational data on the natural course and outcome are crucial for achieving accurate prognostication. Evaluation of biosamples from longitudinal cohort studies may help elucidating mechanisms that drive the progression from arthralgia to clinically evident RA and may reveal targets for potential intervention. Treating patients before they present with clinical arthritis will make it impossible to obtain reliable information about the natural course of the disease. We may then end up in the belief that we are treating the correct patients, but without appropriate scientific endorsement. This scenario bears resemblance to the current situation for patients with UA. For now, we should learn lessons from the past and remain reluctant to start treatment in the absence of clinical arthritis.

Contributors Both authors have written the manuscript and approved the final version.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iD

Annette van der Helm-van Mil http://orcid.org/0000-0001-8572-1437

REFERENCES

1 Combe B, Landewe R, Daien CI, *et al*. 2016 update of the EULAR recommendations for the management of early arthritis. *Ann Rheum Dis* 2017;76:948–59.
2 Smolen JS, Landewé R, Bijlsma J, *et al*. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis* 2017;76:960–77.
3 Available: https://www.trialregister.nl/trial/4599
4 EudraCT number: 2013-003413-18.
5 EudraCT number: 2014-000555-93.
6 Available: https://clinicaltrials.gov/ct2/show/NCT02603146
7 van der Helm-van Mil AHM, le Cessie S, van Dongen H, *et al*. A prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: how to guide individual treatment decisions. *Arthritis Rheum* 2007;56:433–40.
8 van der Helm-van Mil AHM, Detert J, le Cessie S, *et al*. Validation of a prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: moving toward individualized treatment decision-making. *Arthritis Rheum* 2008;58:2241–7.
9 McNally E, Keogh C, Galvin R, *et al*. Diagnostic accuracy of a clinical prediction rule (CPR) for identifying patients with recent-onset undifferentiated arthritis who are at a high risk of developing rheumatoid arthritis: a systematic review and meta-analysis. *Semin Arthritis Rheum* 2014;43:498–507.
10 Kuriya B, Cheng CK, Chen HM, *et al*. Validation of a prediction rule for development of rheumatoid arthritis in patients with early undifferentiated arthritis. *Ann Rheum Dis* 2009;68:1482–5.
11 Ghosh K, Chatterjee A, Ghosh S, *et al*. Validation of Leiden score in predicting progression of rheumatoid arthritis in undifferentiated arthritis in Indian population. *Ann Med Health Sci Res* 2016;6:205–10.
12 Arana-Guajardo A, Pérez-Barbosa L, Vega-Morales D, *et al*. Application of a prediction model for the progression of rheumatoid arthritis in patients with undifferentiated arthritis. *Reumatol Clin* 2014;10:360–3.
13 Aletaha D, Neogi T, Silman AJ, *et al*. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League against rheumatism collaborative initiative. *Ann Rheum Dis* 2010;69:1580–8.
14 van Dongen H, van Aken J, Lard LR, *et al*. Efficacy of methotrexate treatment in patients with probable rheumatoid arthritis: a double-blind, randomized, placebo-controlled trial. *Arthritis Rheum* 2007;56:1424–32.
15 Saleem B, Mackie S, Quinn M, *et al*. Does the use of tumour necrosis factor antagonist therapy in poor prognosis, undifferentiated arthritis prevent progression to rheumatoid arthritis? *Ann Rheum Dis* 2008;67:1178–80.

16 Verstappen SMM, McCoy MJ, Roberts C, *et al*. Beneficial effects of a 3-week course of intramuscular glucocorticoid injections in patients with very early inflammatory polyarthritis: results of the STIVEA trial. *Ann Rheum Dis* 2010;69:503–9.
17 Machold KP, Landewé R, Smolen JS, *et al*. The stop arthritis very early (save) trial, an international multicentre, randomised, double-blind, placebo-controlled trial on glucocorticoids in very early arthritis. *Ann Rheum Dis* 2010;69:495–502.
18 Emery P, Durez P, Dougados M, *et al*. Impact of T-cell costimulation modulation in patients with undifferentiated inflammatory arthritis or very early rheumatoid arthritis: a clinical and imaging study of abatacept (the adjust trial). *Ann Rheum Dis* 2010;69:510–6.
19 Burgers LE, Allaart CF, Huizinga TWJ, *et al*. Brief report: clinical trials aiming to prevent rheumatoid arthritis cannot detect prevention without adequate risk stratification: a trial of methotrexate versus placebo in undifferentiated arthritis as an example. *Arthritis Rheumatol* 2017;69:926–31.
20 Hilliquin S, Hugues B, Mitrovic S, *et al*. Ability of disease-modifying antirheumatic drugs to prevent or delay rheumatoid arthritis onset: a systematic literature review and meta-analysis. *Ann Rheum Dis* 2018;77:1099–106.
21 van Tuyl LHD, Sadlonova M, Hewlett S, *et al*. The patient perspective on absence of disease activity in rheumatoid arthritis: a survey to identify key domains of patient-perceived remission. *Ann Rheum Dis* 2017;76:855–61.
22 Ajeganova S, van Steenbergen HW, van Nies JAB, *et al*. Disease-modifying antirheumatic drug-free sustained remission in rheumatoid arthritis: an increasingly achievable outcome with subsidence of disease symptoms. *Ann Rheum Dis* 2016;75:867–73.
23 van Steenbergen HW, Aletaha D, Beaart-van de Voorde LJJ, *et al*. EULAR definition of arthralgia suspicious for progression to rheumatoid arthritis. *Ann Rheum Dis* 2017;76:491–6.
24 Burgers LE, Siljehult F, Ten Brinck RM, *et al*. Validation of the EULAR definition of arthralgia suspicious for progression to rheumatoid arthritis. *Rheumatology* 2017;56:2123–8.
25 Rakieh C, Nam JL, Hunt L, *et al*. Predicting the development of clinical arthritis in anti-CCP positive individuals with non-specific musculoskeletal symptoms: a prospective observational cohort study. *Ann Rheum Dis* 2015;74:1659–66.
26 Nam JL, Hensor EMA, Hunt L, *et al*. Ultrasound findings predict progression to inflammatory arthritis in anti-CCP antibody-positive patients without clinical synovitis. *Ann Rheum Dis* 2016;75:2060–7.
27 van de Stadt LA, Witte BI, Bos WH, *et al*. A prediction rule for the development of arthritis in seropositive arthralgia patients. *Ann Rheum Dis* 2013;72:1920–6.
28 van Steenbergen HW, Mangnus L, Reijnierse M, *et al*. Clinical factors, anticitrullinated peptide antibodies and MRI-detected subclinical inflammation in relation to progression from clinically suspect arthralgia to arthritis. *Ann Rheum Dis* 2016;75:1824–30.
29 Hensvold A, Kisten Y, Circiumaru A, *et al*. Development of ultrasound detectable arthritis among AcpA positive subjects with musculoskeletal symptoms: the risk RA prospective study. *Arthritis Rheumatol* 2019;71.
30 Hensvold AH, Frisell T, Magnusson PKE, *et al*. How well do ACPA discriminate and predict RA in the general population: a study based on 12 590 population-representative Swedish twins. *Ann Rheum Dis* 2017;76:119–25.
31 Haji Y, Rokutanda R, Kishimoto M, *et al*. Anti-cyclic citrullinated protein antibody (ACPA) positivity in general population and follow–up results for ACPA positive persons [abstract]. *Arthritis Rheumatol* 2016;68:512.
32 Ramos-Remus C, Castillo-Ortiz JD, Aguilar-Lozano L, *et al*. Autoantibodies in prediction of the development of rheumatoid arthritis among healthy relatives of patients with the disease. *Arthritis Rheumatol* 2015;67:2837–44.
33 Falahee M, Finckh A, Raza K, *et al*. Preferences of patients and at-risk individuals for preventive approaches to rheumatoid arthritis. *Clin Ther* 2019;41:1346–54.
34 Mosor E, Stoffer-Marx M, Steiner G, *et al*. *I would never take preventive medication!* Perspectives and information needs of people who underwent predictive tests for rheumatoid arthritis. *Arthritis Care Res* 2019.

CLINICAL SCIENCE

Efficacy and safety of NI-0101, an anti-toll-like receptor 4 monoclonal antibody, in patients with rheumatoid arthritis after inadequate response to methotrexate: a phase II study

Emmanuel Monnet ,[1] Ernest H Choy,[2] Iain McInnes,[3] Tamta Kobakhidze,[4] Kathy de Graaf,[1] Philippe Jacqmin,[5] Geneviève Lapeyre,[1] Cristina de Min[1]

Handling editor Josef S Smolen

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-216487).

[1]Clinical Development, NovImmune SA, Plan les Ouates, Geneva, Switzerland
[2]Institute of Infection and Immunity, Cardiff University School of Medicine, Cardiff, UK
[3]Institute of Infection, Immunity and Inflammation, University of Glasgow, Glasgow, UK
[4]University Clinic, High Technology Medical Center, Tbilisi, Georgia
[5]MsN, SPRL, Dinant, Belgium

Correspondence to
Dr Emmanuel Monnet, Clinical Development, NovImmune SA, Geneva, Plan les Ouates 1228, Switzerland;
emonnet@novimmune.com

EHC and IM contributed equally.

Received 16 October 2019
Revised 3 December 2019
Accepted 13 December 2019
Published Online First
31 December 2019

To cite: Monnet E, Choy EH, McInnes I, *et al*. *Ann Rheum Dis* 2020;**79**:316–323.

ABSTRACT

Objectives Anti-citrullinated protein antibodies (ACPAs) form immune complexes with citrullinated proteins binding toll-like receptor (TLR) 4, which has been proposed as a mediator of rheumatoid arthritis (RA). NI-0101 is a first-in-class humanised monoclonal antibody blocking TLR4, as confirmed by inhibition of in vivo lipopolysaccharide-induced cytokine release in healthy volunteers. This study was design to confirm preclinical investigations supporting a biomarker-driven approach for treatment of patients with RA who present positive for these immune complexes.

Methods Placebo-controlled, double-blind, randomised (2:1) trial of the tolerability and efficacy of NI-0101 (5 mg/kg, every 2 weeks for 12 weeks) versus placebo in ACPA-positive RA patients with inadequate response to methotrexate. Efficacy measures included Disease Activity Score (28-joint count) with C reactive protein (DAS28-CRP), European League Against Rheumatism (EULAR) good and moderate responses, and American College of Rheumatology (ACR) 20, ACR50 and ACR70 responses. Subgroup analyses defined on biomarkers were conducted. Pharmacokinetics, pharmacodynamics and safety were reported.

Results 90 patients were randomised (NI-0101 (61) and placebo (29)); 86 completed the study. No significant between-group difference was observed for any of the efficacy endpoints. Subgroup analyses using baseline parameters as covariants did not reveal any population responding to NI-0101. Treatment-emergent adverse events occurred in 51.7% of patients who received placebo versus 52.5% for NI-0101.

Conclusions We demonstrate for the first time that in RA, a human immune-mediated inflammatory disease, blocking the TLR4 pathway alone does not improve disease parameters. Successful targeting of innate immune pathways in RA may require broader and/or earlier inhibitory approaches.

Key messages

What is already known about this subject?

- Citrullinated proteins and anti-citrullinated protein antibodies forming immune complexes belong to the damage-associated molecular pattern family, participating in innate immunity and are expressed in inflammatory conditions, such as in rheumatoid arthritis (RA).
- Immune and stromal cells are activated by these immune complexes via cellular receptors, including toll-like receptor (TLR) 4. NI-0101 is a humanised immunoglobulin G1κ monoclonal antibody engineered to bind to and block the activation of human TLR4, which has demonstrated a predictable pharmacokinetics, good safety profile and inhibition of in vivo lipopolysaccharide-induced cytokine production in healthy volunteers.

What does this study add?

- We assessed for the first time, in a placebo-controlled, double-blind, randomised study, the tolerability and efficacy of TLR4 blockade in RA patients with inadequate response to methotrexate (MTX). Study results indicated no significant differences between treatment arms for any of the clinical efficacy and pharmacodynamics endpoints included in prespecified subgroups positive for antibodies against selected citrullinated proteins.

How might this impact on clinical practice or future developments?

- This study demonstrated that the blockage of TLR4 is likely not a relevant target in RA patients with inadequate response to MTX and established disease, its role remains to be determined.
- Successful targeting of innate immune pathways in RA, and potentially also in other chronic inflammatory diseases, may require broader or earlier inhibitory approaches.

INTRODUCTION

Both innate and adaptive immune pathways are implicated in the pathogenesis of rheumatoid arthritis (RA).[1] Anti-citrullinated protein antibodies (ACPAs) are characteristic of RA and may be present prior to the emergence of clinical symptoms of the disease.[2 3] Citrullinated proteins and ACPAs form immune complexes[4 5] which belong to the damage-associated molecular pattern (DAMP) family.[6] DAMPs are important regulators of innate inflammatory responses. They drive pathogenic processes in RA by activating both immune and stromal cells by stimulating cellular receptors, including

toll-like receptor (TLR) 4.[7 8] This pattern recognition receptor can be activated by immune complexes formed by citrullinated proteins, including matrix-derived molecules (eg, citrullinated-fibrinogen) and their associated autoantibodies (ACPAs).[9–13] These molecules are upregulated in some patients with RA and are expressed in the synovium.[14] Numerous preclinical mechanistic studies have shown the potential role for TLR4 and its ligands in RA.[15–24]

Biological agents currently approved for the treatment of RA block the actions of tumour necrosis factor (TNF)-α or interleukin (IL)-6 receptor, directly interfere with the actions of T cells or deplete B cells.[25] T cell inhibition by abatacept and cytokine signalling reduction by Janus kinase inhibitors have also demonstrated efficacy for the treatment of RA.[26] Numerous targeted therapies are available, but unmet needs in the management of RA remain. Partial and loss of response are common and drug-free remission cannot be achieved in most patients.[27] Moreover, patients who fail one biological agent may receive even less benefit when switching to a second agent, even with a different mechanism of action.[28] This may in part reflect accrual of irreversible articular damage mediating chronicity in synovial pathology.[28] Some patients ultimately become resistant to all currently available therapeutics—so-called difficult-to-treat RA,[29] requiring new therapeutic solutions. Given the evidence supporting a role for TLR4 in RA pathogenesis, we explored inhibition of this pathway as a potential treatment target.

NI-0101 is a humanised immunoglobulin (Ig) G1κ monoclonal antibody engineered to bind to and block the activation of human TLR4. It interferes with TLR4 dimersation, preventing signal transduction through the TLR4 cytoplasmic pathway.[30] It has been demonstrated to inhibit the effects of lipopolysaccharide (LPS) administered to healthy volunteers, which is dependent on FcγRII.[31] The results from in vitro studies have demonstrated a correlation between levels of TLR4 ligands and blockade of innate inflammatory responses by NI-0101.[9]

METHODS

Study design

This was a phase II, proof-of-concept, randomised (2:1), placebo-controlled, double blind, international multicentre study in patients with moderate-to-severe ACPA-positive RA that previously responded inadequately to methotrexate (MTX). Patients received addition of NI-0101 (5 mg/kg administered every 2 weeks for 12 weeks) or placebo to ongoing MTX treatment for 12 weeks. Patients in both treatment arms were stratified on the basis of FcγRIIa genotype (RR/RH and HH) and C reactive protein (CRP) level (above and below 0.7 mg/dL, with a maximum of 25% below 0.7 mg/dL). Patients were followed up for 12 weeks after NI-0101 was stopped.

Patients

Male and female patients ≥18 years old and with body mass indices <30 and >18 kg/m^2 with a diagnosis of RA according to 2010 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) criteria, ACPA positive and disease duration ≥6 months since formal diagnosis were eligible for enrolment. Patients had active RA at screening, characterised by ≥6 of 66 swollen joints and ≥6 of 68 tender joints, confirmed synovitis in ≥1 of the six swollen joints, CRP >0.7 mg/dL or CRP level between 0.3 and 0.7 mg/dL if erythrocyte sedimentation rate (ESR) ≥30 mm/hour, and to have been receiving MTX for ≥3 months and a stable dose/regimen for ≥6 weeks prior to screening.

Patient participation was excluded by a history of autoimmune disease other than RA, prior receipt of a cytotoxic agent other than MTX or immunosuppressive drugs ≤3 months prior to screening (see online supplementary data for more details).

Patient and public involvement

It was not appropriate or possible to involve patients or the public in the design, or conduct, or reporting or dissemination of our research.

Assessments

Efficacy

Efficacy measures included OMERACT RA core outcome set and clinical study reported according to EULAR recommendations on conducting/reporting of clinical trials. Efficacy measures included mean values and changes from baseline in Disease Activity Score including 28-joint count using CRP or ESR (DAS28-CRP, DAS28-ESR); Simplified Disease Activity Index (SDAI) and Clinical Disease Activity Index (CDAI) scores; and proportions of patients achieving EULAR good, moderate and no response; or ACR20, ACR50 and ACR70 responses. Subgroup analyses included assessment of the effects of baseline (study day 0 prior to first treatment administration) patient characteristics and biomarkers (APCA, citrullinated peptide-specific APCA, circulating TLR4 ligands, rheumatoid factor (RF)) on clinical outcomes.

Pharmacokinetics and pharmacodynamics

NI-0101 concentrations was measured preinfusion, throughout the treatment and until the end of the follow-up period. Changes from baseline in CRP, IL-6, IL-1β, IL-8, TNF-α and C-X-C motif chemokine 10 (CXCL10) were evaluated.

Safety

Safety assessments consisted of recording of adverse events (AEs), clinical laboratory values and vital signs; and testing for the presence of antidrug antibodies (ADAs).

Statistical analysis

Study populations included the intent-to-treat-completer (c-ITT) analysis set, defined as all patients who were randomised and completed the treatment period; the per-protocol (PP) analysis set, defined as all patients in the c-ITT population without any major protocol deviations; and the safety (SAF) analysis set, defined as all patients who received at least part of the first infusion of NI-0101 or placebo. Patients were analysed according to the actual treatment received.

Efficacy endpoints were analysed by statistical models including treatment, score for each measure at baseline and randomisation stratification factors (FcγRIIa genotype and CRP level at baseline) as fixed effect covariates. Other covariates, including country, duration of RA, use of non-steroidal anti-inflammatory drugs and glucocorticoids at baseline, baseline joint counts, ESR values, VECTRA DA scores and ACPA level could also be investigated in analyses of DAS28-CRP and ACR50 results.

Calculation of sample size for the randomised treatment arms was based on the change in DAS28-CRP between the NI-0101 and the placebo groups for RR/RH population at week 12 compared with predose. It was estimated that 54 RR/RH patients (NI-0101:placebo; 36:18) gave a power of 80% at a two-sided significance level of 5% assuming a difference in DAS28-CRP of 1 point (SD=1.2) at 12 weeks between treatment and placebo (2:1 ratio). Considering that the population includes ≥66% of

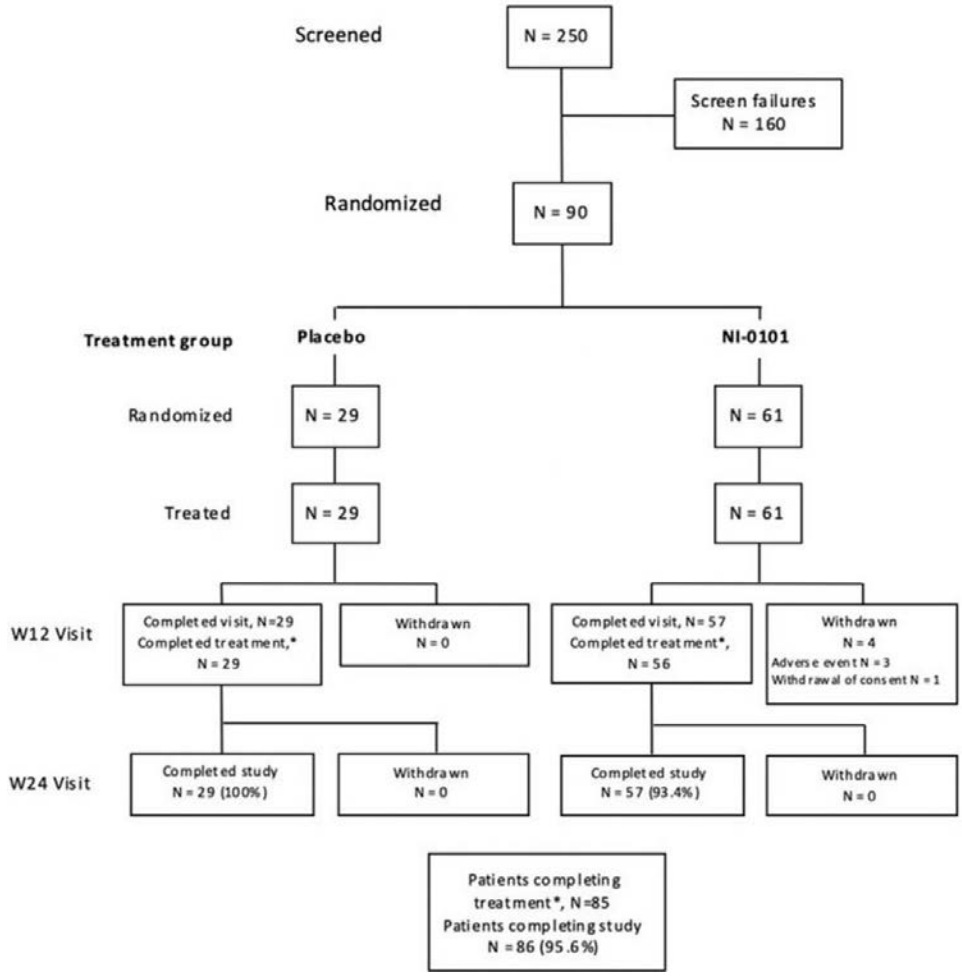


Figure 1 Patient disposition. Data in boxes represent numbers of patients. *Defined as patients who received at least five of the six scheduled infusions and had at least one evaluable efficacy data at week 12.

RR/RH, the total number of patients required to complete the treatment was calculated to be 81 (NI-0101:placebo; 54:27) to ensure at least 54 RR/RH patients completed treatment. Ninety patients were randomised to compensate for dropouts.

RESULTS

Patients and screening phase

Of 250 patients screened for eligibility, 90 were randomised (61 to NI-0101 and 29 to placebo group). All randomised patients received at least part of the first infusion of NI-0101 and 57 completed the week 12 visit along with 29 patients treated with placebo, all of these patients completed the follow-up phase to week 24 (figure 1). Baseline demographic and disease characteristics are summarised in table 1. There were no major imbalances between groups for most individual disease parameters. However, patients in the NI-0101 group had a longer duration of RA (8.5 years vs 5.4 years for placebo) and were younger at the time of RA diagnosis (45.7 years vs 51.2 years for placebo). The mean CRP level was also higher for patients allocated to receive NI-0101 (18.3 mg/L vs 13.4 mg/L for placebo) at baseline, whereas CRP levels at screening were slightly higher in the placebo group. CRP levels decreased between screening and baseline for most patients in each group, but the decline was greater for those who received placebo. Post hoc analysis demonstrated that the magnitude of the CRP decrease was dependent on the recruitment site of origin.

Efficacy

Both treatment groups demonstrated similar decreases from baseline to week 12 in DAS28-CRP with no significant between-group difference (figure 2A); a similar pattern was observed for DAS28-ESR (figure 2B). CDAI and SDAI scores decreased by approximately 40% from baseline to week 12, again with no significant differences between treatment groups (figure 2C,D). The proportion of patients achieving EULAR responses (good or moderate) increased with treatment. By week 12, 27.6% and 26.0% of patients in the placebo and NI-0101 groups, respectively, had achieved EULAR good responses; and 55.2% and 53.6% had achieved EULAR moderate responses (figure 3A). There were no significant between-group differences in ACR responses at week 12; 55.2% and 58.9% of patients in the placebo and NI-0101 groups, respectively, achieved ACR20 responses; 20.7% and 14.3% achieved ACR50 responses, and 10.3% and 10.7% achieved ACR70 responses (figure 3B–D). Swollen and tender joint counts also declined from baseline in both treatment groups. The changes in swollen joints from baseline to week 12 were –6.1 and –7.1 for the placebo and NI-0101 groups, respectively; and the respective values for tender joints were –6.3 and –8.1.

Subgroup analysis indicated no significant effects on stratification by CRP and FcγRIIa genotype for DAS28-CRP or ACR50 response. All subgroup analyses, based on levels of prespecified biomarkers (ACPA, RF, cFb-IC, anti-citrullinated protein/

Table 1 Baseline demographic and clinical characteristics

Baseline characteristic	Measure	Placebo, n (%) (n=29)	NI-0101, n (%) (n=61)
Sex, n (%)	Males	6 (20.7)	11 (18.0)
	Females	23 (79.3)	50 (82.0)
Race, n (%)	White	29 (100.0)	61 (100.0)
Age (years)	Mean (SD)	57.1 (13.07)	54.6 (11.10)
	Median (range)	59.1 (20–79)	56.3 (23–76)
Weight (kg)	Mean (SD)	68.8 (15.46)	71.4 (13.30)
	Median (range)	66.5 (47.0–103.9)	70.8 (45.6–98.9)
BMI (kg/m²)	Mean (SD)	25.2 (4.01)	26.3 (3.43)
	Median (range)	25.9 (18.0–29.8)	26.3 (18.4–32.0)
Duration of RA	Mean years (SD)	5.4 (4.82)	8.5 (7.86)
	Range	0.5–17.1	0.5–33.1
Age at RA diagnosis	Mean years (SD)	51.2 (13.62)	45.7 (11.56)
	Range	18–69	21–67
Steroid dose category	No steroid given	9 (31.0)	20 (32.8)
	1–5 mg	8 (27.6)	6 (9.8)
	5–10 mg	12 (41.4)	35 (57.4)
MTX dose category (mg/week)	3.5–10 mg	2 (6.9)	2 (3.3)
	10–20 mg	25 (86.2)	55 (90.2)
	20–25 mg	2 (6.9)	4 (6.6)
CRP (mg/L)	Mean (SD)	13.4 (14.03)	18.3 (26.63)
ESR (mm/hour)	Mean (SD)	43.1 (16.51)	45.3 (24.26)
RF (IU/mL)	Mean (SD)	127.6 (146.36)	149.3 (175.72)
ACPA (U/mL)	Mean (SD)	962.6 (1730.87)	676.2 (1072.80)
DAS28-CRP	Mean (SD)	5.8 (0.82)	5.9 (0.94)
DAS28-ESR	Mean (SD)	6.6 (0.88)	6.6 (0.91)
68-tender joint counts	Mean (SD)	28.9 (14.07)	27.5 (15.89)
66-swollen joint counts	Mean (SD)	16.3 (7.92)	16.8 (8.96)

ACPA, anti-citrullinated protein antibody; BMI, body mass index; CRP, C reactive protein; DAS28, Disease Activity Score, including a 28-joint count; ESR, erythrocyte sedimentation rate; MTX, methotrexate; RA, rheumatoid arthritis; RF, rheumatoid factor.

peptide antibodies, TLR4 ligands) measured at baseline and post hoc analyses using baseline disease-related parameters failed to demonstrate any significant treatment effects in any of the subgroups.

Pharmacokinetics

The NI-0101 pharmacokinetics (PK) profile showed expected concentrations with an elimination was consistent with simulations. Throughout the treatment period, NI-0101 concentrations were maintained above the targeted threshold of 10 000 ng/mL in the majority of patients. The half-life for the linear elimination phase was estimated to be approximately 6.4 days.

Pharmacodynamics

There were no significant differences between treatment groups for all biomarkers evaluated (table 2). Analysis of changes in CRP levels from baseline to week 12 showed small increases for both treatment groups (see online supplementary data).

Safety

NI-0101 infusions every 2 weeks elicited an acceptable safety and tolerability profile in patients with RA. The Data Monitoring Committee did not request for changes in the conduct of the study and no deaths were reported. Treatment-emergent adverse events (TEAEs) reported from baseline to week 24 occurred in similar proportions of patients in the placebo and NI-0101 groups; 51.7% and 52.5%, respectively (table 3). Five patients (5.6%) reported TEAEs considered to be related to NI-0101. One patient in the placebo group and three patients in the NI-0101 group discontinued treatment due to TEAEs; however, only one of these TEAEs (an infusion-related reaction (IRR)) was assessed as having a relationship with the administration of NI-0101. One patient in the placebo group experienced a serious adverse event (AE) (appendicitis and peritoneal abscess) as did three patients in the NI-0101 group (severe IRR, diagnosis of adenocarcinoma of the colon and diagnosis of ovarian cancer). In three other patients of the NI-0101 group, non-serious events (mild dermatitis, moderate urinary tract infection and alanine

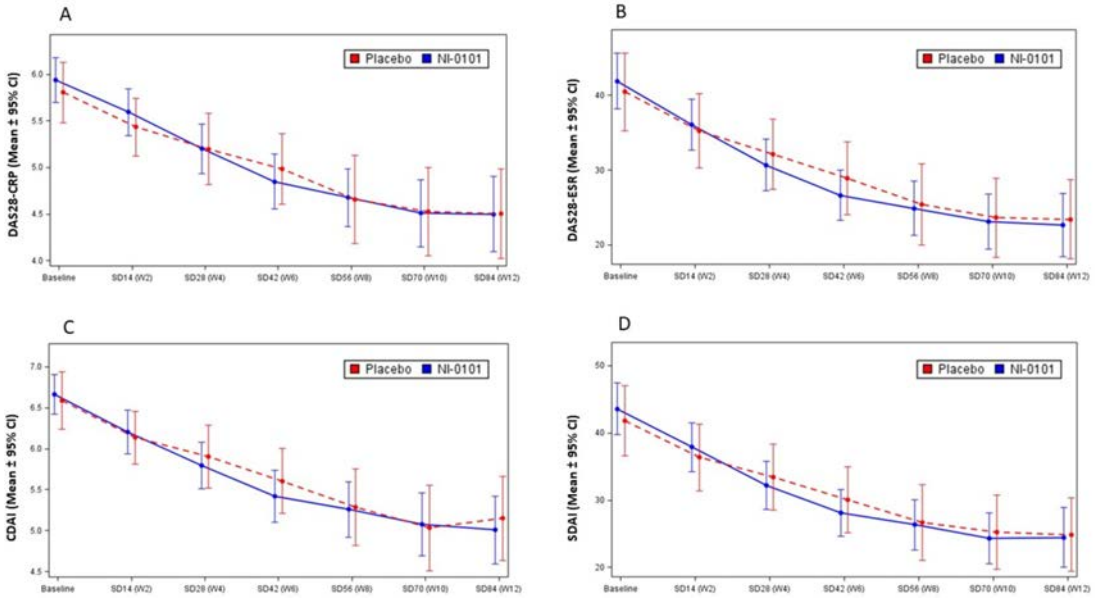


Figure 2 (A) DAS28-CRP scores. (B) DAS28-ESR scores. (C) CDAI scores. (D) SDAI scores. All values are means±95% CI. Placebo, n=28; NI-0101, n=54. CDAI, Clinical Disease Activity Index ; DAS28-CRP, Disease Activity Score (28-joint count) with C reactive protein; AS28-ESR, Disease Activity Score (28-joint count) with erythrocyte sedimentation rate; SDAI, Simplified Disease Activity Index.

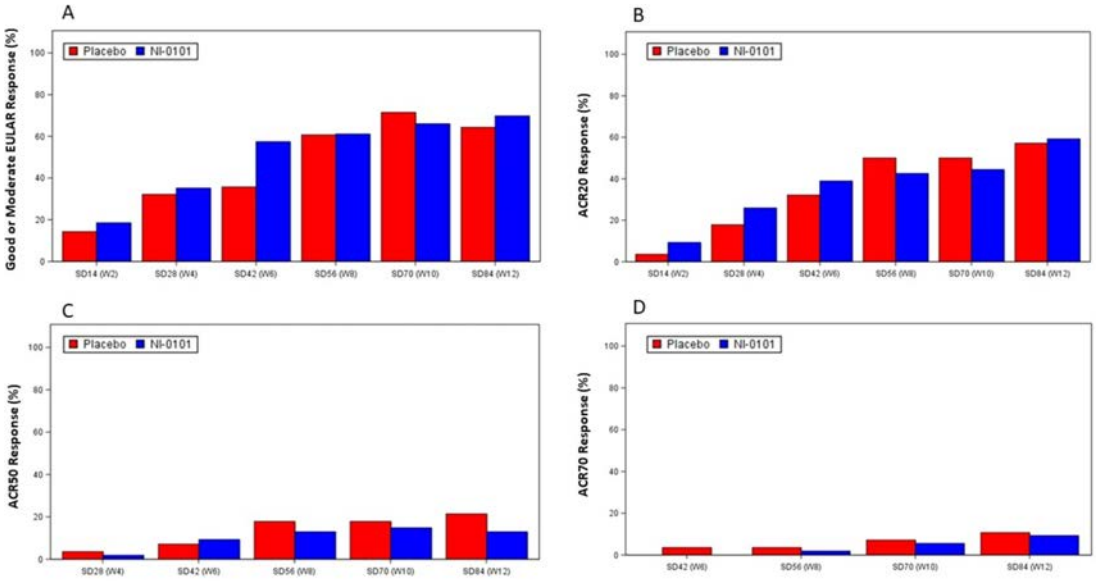


Figure 3 (A) Percentage of patients achieving EULAR good or moderate responses. (B–D) Percentages of patients achieving ACR20, 50 and 70 responses. ACR, American College of Rheumatology; EULAR, European League Against Rheumatism. Placebo, n=28; NI-0101, n=54. EULAR response at week 12: or 1.36, 95% CI (0.51; 3.67), p value 0.5381. ACR20 response at week 12: OR 1.07, 95% CI (0.42; 2.72), p value 0.8948. ACR50 response at week 12: OR 0.63, 95% CI (0.18; 2.18), p value 0.4665. ACR70 response at week 12: OR 0.94, 95% CI (0.20; 4.32), p value 0.9318.

aminotransferase grade 2 increase) were reported as related to NI-0101 but did not result in treatment discontinuation.

Infections were the most frequently reported AEs (11.5% and 13.8% in the NI-0101 and placebo groups, respectively). None of the infections reported in the NI-0101 group were rated as severe or serious. Most were respiratory tract infections commonly observed during autumn and winter. All were mild or moderate in intensity. Infections were not considered related to study treatment, except one moderate urinary tract infection.

No safety signals were identified for other safety parameters.

DISCUSSION

This is the first study to assess the efficacy of TLR4 inhibition in patients with RA or indeed with an immune-mediated inflammatory disease. The efficacy analysis showed consistent, but moderate, improvements for all endpoints evaluated for both treatment groups but no significant differences between addition of NI-0101 or placebo to MTX. Response level observed in the placebo group was higher than typically reported for clinical studies in this population, particularly for moderate response measured either by EULAR criteria or by ACR20 response. Good EULAR responses and achievement of ACR50 and ACR70 improvements in the placebo group were closer to values reported previously for patients with inadequate responses to MTX and continued on this treatment, although on the high end of such response rates.[32 33] In general, the NI-0101 treatment group showed similar or worse responses than the placebo group at week 12. Moreover, the improvements noted were lower than observed when other targeted DMARDs (biologics or small molecules) have been added to therapy in MTX-IR patients with RA.[34 35] Despite clinical improvement in both treatment groups, there was no significant reduction from baseline

Table 2 Assessments of inflammatory markers

		Change from baseline to W12, mean (SD)		P value	
Parameter, pg/mL	Baseline value, all patients, mean (SD)	Placebo (n=28)	Ni-0101 (n=54)	Treatment effect	Baseline value effect
CRP	15.6 (17.27)	−0.3 (2.83)	0.6 (2.11)	0.7688	–
IL-6	19.3 (59.2)	−5.3 (38.04)	−2.4 (18.22)	0.3978	<0.0001
GM-CSF*	9.4 (0)	0 (0)	0 (0)	–	–
IL-17A*	15.4 (0)	0 (0)	0 (0)	–	–
IL-10	0.8 (0.98)	0 (0.66)	0.3 (2.41)	0.5148	0.0319
IL-1β	1.2 (0.06)	0 (0)	0.1 (0.58)	–	<0.0001
IL-8	23.7 (18.87)	0.3 (12.24)	−3.0 (15.73)	0.2698	<0.0001
INF-γ	15.5 (30.05)	7.5 (31.50)	−0.2 (40.57)	0.7860	<0.0001
TNF-α	5.6 (11.99)	2.0 (11.49)	−0.1 (1.85)	0.5548	<0.0001
CXCL10	651.9 (542.8)	−17.4 (506.73)	−35.7 (338.77)	0.5624	<0.0001
MCP-1	422.9 (162.18)	13.4 (127.29)	−18.9 (124.58)	0.2667	0.0027

'Baseline value effect' assesses the effect of variability at baseline on the tested outcome. Here, baseline variability reported for the measured cytokines is higher than the tested treatment effect.

*Values were below limit of quantification.

CRP, C-reactive protein; CXCL10, C-X-C motif chemokine 10;GM-CSF, granulocyte-macrophage colony-stimulating factor; IL, interleukin; INF, interferon;IP-10, interferon gamma-induced protein 10; MCP, monocyte chemoattractant protein;TNF, tumour necrosis factor; W, week.

Table 3 TEAEs through 24 weeks

	Placebo, n (%) (n=29)	NI-0101, n (%) (n=61)
Pretreatment AEs	1 (3.4)	2 (3.3)
TEAEs to week 24	15 (51.7)	32 (52.5)
TEAEs related to administered treatment	0	5 (8.2)
Serious TEAEs	1 (3.4)	3 (4.9)
TEAEs leading to treatment discontinuation	1 (3.4)	3 (4.9)
TEAEs leading to death	0	0
TEAEs related to potential IRRs	3 (10.3)	9 (14.8)
TEAEs related to infections	5 (17.2)	17 (27.9)
TEAEs by highest severity		
Mild	6 (20.7)	12 (19.7)
Moderate	9 (31.0)	17 (27.9)
Severe	0	3 (4.9)
Life threatening	0	0
Fatal	0	0
Missing	0	0
TEAEs experienced by ≥5% of patients in either treatment group		
Nasopharyngitis	2 (10.3)	3 (4.9)
Upper respiratory tract infection	1 (3.4)	4 (6.6)
Condition aggravated	0	5 (8.2)

IRRs, infusion-related reactions; TEAE, treatment emergent adverse event.

in CRP, an objective measure of inflammation, for patients receiving either placebo or NI-0101 added to MTX. A potential therapeutic response to MTX background therapy during screening was observed based on CRP decrease, possibly driven by higher adherence to background treatment between screening and randomisation.

The absence of a significant effect of adding NI-0101 to MTX was further confirmed by the lack of treatment-associated changes in levels of cytokines downstream from TLR4 and known to be involved in the inflammation characteristic of RA.[36] The lack of effect of NI-0101 versus placebo on levels of inflammatory molecules evaluated in this study extended to IL-6, TNF-α, IL-8 and IL-1β, all of which have been shown to be elevated in monocytes from synovial fluid through TLR4 signalling and blocked by exposure to NI-0101 in vitro.[9 37]

During the follow-up period, when the patient and treating physician knew that NI-0101 was no longer being administered (while remaining blinded to prior treatment allocation), the results for all efficacy endpoints remained stable or decreased by similar amounts in both treatment arms. As the elimination half-life of NI-0101 is 6.4 days, it would have been reasonable to expect some continued benefit after treatment withdrawal, if it had significant efficacy.

Preplanned subgroup analyses using baseline levels of TLR4-related biomarkers were conducted to test the hypothesis that RA patients with elevated levels of TLR4 ligands (eg, citrullinated protein immune complexes) would have an increased response to the addition of NI-0101 to MTX. However, patient segmentation on the basis of the selected biomarkers failed to demonstrate any benefit of NI-0101 versus placebo. Furthermore, post hoc subgroup analyses using baseline disease and demographic parameters, including, but not limited to, baseline CRP levels and variations during screening, country of origin and disease duration, were conducted to potentially identify confounding parameters, but none showed a statistically significant effect on any between-treatment differences. The PK results from this study and PK/pharmacodynamic analysis from a prior study[31] suggest that the levels of NI-0101 achieved in the patients in this trial were sufficient to achieve TLR4 pathway blockade between two dosing intervals, regardless of the FcγRIIa polymorphism. Thus, it is unlikely that insufficient levels of NI-0101 contributed to the observed lack of clinical effect.

Given that NI-0101 has been shown to be a potent inhibitor of TLR4, as demonstrated by the lack of induction of inflammatory cytokines after in vivo LPS administration in healthy volunteers after having received NI-0101 and that literature on pathogenic processes in RA reports the involvement of the stimulation of this receptor,[7–12 31] the lack of significant clinical and pharmacodynamic effects in this study are surprising. It is possible that redundancy in TLR signalling may underlie the lack of effect of TLR4 blockade in this trial. In fact, TLR2, TLR4, TLR5 and TLR7 have all been considered to be potentially involved in the pathology of RA.[38] It cannot be excluded that NI-0101 may provide clinical benefit when combined with other targeted agents. Indeed, the preclinical hypothesis tested in this study was supported by the observed correlation in vitro between NI-0101 response and the presence of specific immune complexes against citrullinated proteins.[9] The presence of antibodies against citrullinated proteins has been reported even before the first clinical manifestation of RA. It is conceivable, perhaps that immune complexes signalling through TLR4 could play a significant pathogenic role in early RA, whereas other inflammatory processes are predominant when RA is already established and therefore blocking TLR4 may not provide any benefit.

We demonstrate satisfactory safety and tolerability of TLR4 inhibition with NI-0101. There were no significant differences between treatment groups in the frequency of AEs. The type and intensity of AEs reported in this study were similar to those observed in prior clinical trials in similar patient cohorts,[39 40] and of the three serious AEs (IRR, adenocarcinoma of the colon and ovarian cancer) reported in the NI-0101 group, only the IRR was related to NI-0101 administration.

TLR4 has been shown to play an important role in immune response to Gram-negative bacteria.[37] However, the results suggest no increased risk for infections with NI-0101 and are consistent with findings from healthy volunteers who received NI-0101, as well as those obtained with other molecules targeting the same pathways.[31 41 42] No systemic Gram-negative infections were reported. The incidence of urinary tract infections (6.6%), all in female patients, appeared no greater than that reported for postmenopausal women who constitute the majority of the RA population.[43 44]

This study demonstrated that the blockage of TLR4 is likely not a relevant target in RA patients with inadequate response to MTX, as shown by the absence of NI-0101 effect versus placebo on clinical endpoints or on changes in levels of inflammatory cytokines or chemokines. In addition, none of the subgroup analyses identified a subset of patients that received benefit from NI-0101. The results showed an expected PK profile, desired concentrations and no safety concerns for NI-0101. The lack of significant effect of NI-0101 in this well-controlled prospective clinical trial indicates that blocking the TLR4 pathway alone is unlikely to benefit patients with established RA. The role of TLR4 and of anti-citrullinated antibodies forming immune complexes in prior diagnosis and in early RA remains to be established. The good NI-0101 safety and PK profiles support further exploration in other diseases, in particular when microbial products are involved in inflammatory diseases or when high microbial translocation is observed (eg, HIV).

Acknowledgements Novimmune and the authors thank the patients, study sites and investigators who participated in this clinical study. The authors thank Novimmune clinical operation (Hannah Frederiksen, Cristina Darie-Chrilla, MarthaLee Reynolds), drug safety (Yuanyuan Lu), data management (Nathalie Brisset, Jan Richardet) and bioanalytical (Anna Ahomaa, Maxime Legues, Stéphanie Teixeira, Joris Venet, Adeline Besson, Sarah Collomb, Estelle Cochet, Lydie Meynet, Clotilde Cettour, Amandine Berthet, Robert Nelson, Laetitia Sordé) team members who supported the study conduct, ArgintINT who performed the monitoring, Cytel who provided data management and statistics services, and Richard Kay who provided strategic statistical support. Medical writing support was provided by Julia Powell, Robert Rhoades and Florence Manger. Individual anonymised participant data, pooled study data, protocol or statistical analysis plan will not be shared unless a specific data sharing agreement is established with Light Chain Bioscience. For inquiries about data sharing, please send request at NI-0101@lightchainbio.com.

Contributors EM, EHC, IMcI, KdG, PJ, GL and CdM participated in the design of the study. EM, KdG, GL and TK participated in data collection. EM, KdG and GL participated in data analysis. EM, EHC, IMcI, KdG, PJ, GL and CdM participated in interpreting the data, in writing and in critically reviewing the manuscript. All authors approved the final version. EHC and IMcI contributed equally to the study design and data interpretation.

Funding Novimmune SA was the sponsor of the study.

Competing interests EM, KdG, GL and CdM were employees and stock options holders of Novimmune SA. PJ was consultant of Novimmune SA. EHC and IMcI were consultants of Novimmune SA: EHC received consultancy fees or grants from UCB, Pfizer, BioCancer, Biogen, Novartis, Roche, Amgen, Chugai, Eli Lilly, Sanofi, Abbvie, Janssen, Gilead and Bristol Myer Squibbs. IMcI received consultancy fees and grants from Celgen, Janssen, Novartis, Beorhinger Ingelheim, Abbvie, Eli Lilly, Bristol Myer Squibbs, GlaxoSmithKline and Pfizer. TK received Investigator fees from Novimmune SA to conduct the study. Novimmune SA and Genentech entered into a collaboration agreement for the development of NI-0101, under this agreement Novimmune SA received funding from Genentech.

Patient consent for publication Not required.

Ethics approval All relevant study documentation and amendments were approved by Independent Ethics Committees. The study was conducted in accordance with the principles set forth in the Declaration of Helsinki, the Guidelines of the International Council for Harmonisation (ICH) on Good Clinical Practice (GCP) Guideline E6 (R2) (EMA/CPMP/ICH/135/95), European Union (EU) Directive 95/46/EC, and other applicable regulatory requirements. Patients provided informed written consent prior to any study procedures.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Individual anonymised participant data, pooled study data, protocol or statistical analysis plan will not be shared unless a specific data sharing agreement is established with Light Chain Bioscience. For inquiries about data sharing, please send request at NI-0101@lightchainbio.com.

ORCID iD

Emmanuel Monnet http://orcid.org/0000-0001-5599-8240

REFERENCES

1 McInnes IB, Schett G. The pathogenesis of rheumatoid arthritis. *N Engl J Med Overseas Ed* 2011;365:2205–19.
2 Rantapää-Dahlqvist S, de Jong BAW, Berglin E, *et al*. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. *Arthritis Rheum* 2003;48:2741–9.
3 Nielen MMJ, van Schaardenburg D, Reesink HW, *et al*. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. *Arthritis Rheum* 2004;50:380–6.
4 Kurowska W, Kuca-Warnawin EH, Radzikowska A, *et al*. The role of anti-citrullinated protein antibodies (AcpA) in the pathogenesis of rheumatoid arthritis. *Cejoi* 2017;42:390–8.
5 Sohn DH, Rhodes C, Onuma K, *et al*. Local joint inflammation and histone citrullination in a murine model of the transition from preclinical autoimmunity to inflammatory arthritis. *Arthritis Rheumatol* 2015;67:2877–87.
6 Stolberg-Stolberg JA, Furman BD, Garrigues NW, *et al*. Effects of cartilage impact with and without fracture on chondrocyte viability and the release of inflammatory markers. *J Orthop Res* 2013;31:1283–92.
7 Huang Q-Q, Pope RM. The role of toll-like receptors in rheumatoid arthritis. *Curr Rheumatol Rep* 2009;11:357–64.
8 Goh FG, Midwood KS. Intrinsic danger: activation of Toll-like receptors in rheumatoid arthritis. *Rheumatology* 2012;51:7–23.
9 Hatterer E, Shang L, Simonet P, *et al*. A specific anti-citrullinated protein antibody profile identifies a group of rheumatoid arthritis patients with a toll-like receptor 4-mediated disease. *Arthritis Res Ther* 2016;18.
10 Smiley ST, King JA, Hancock WW. Fibrinogen stimulates macrophage chemokine secretion through toll-like receptor 4. *J Immunol* 2001;167:2887–94.
11 Wenink MH, Santegoets KCM, Roelofs MF, *et al*. The inhibitory Fc gamma IIb receptor dampens TLR4-mediated immune responses and is selectively up-regulated on dendritic cells from rheumatoid arthritis patients with quiescent disease. *J Immunol* 2009;183:4509–20.
12 Anquetil F, Clavel C, Offer G, *et al*. IgM and IgA rheumatoid factors purified from rheumatoid arthritis sera boost the Fc receptor- and complement-dependent effector functions of the disease-specific Anti–Citrullinated protein autoantibodies. *J Immunol* 2015;194:3664–74.
13 Manivel VA, Sohrabian A, Rönnelid J. Granulocyte-augmented chemokine production induced by type II collagen containing immune complexes is mediated via TLR4 in rheumatoid arthritis patients. *Eur J Immunol* 2016;46:2822–34.
14 Ospelt C, Brentano F, Rengel Y, *et al*. Overexpression of toll-like receptors 3 and 4 in synovial tissue from patients with early rheumatoid arthritis: toll-like receptor expression in early and longstanding arthritis. *Arthritis Rheum* 2008;58:3684–92.
15 Achek A, Shah M, Seo JY, *et al*. Linear and rationally designed stapled peptides abrogate TLR4 pathway and relieve inflammatory symptoms in rheumatoid arthritis rat model. *J Med Chem* 2019;62:6495–511.
16 Li D, Zhou Q, Hu G, *et al*. MiRNA-506 inhibits rheumatoid arthritis fibroblast-like synoviocytes proliferation and induces apoptosis by targetting TLR4. *Biosci Rep* 2019;39.
17 Umekita K, Miyauchi S, Nomura H, *et al*. Neutrophil-derived lactoferrin induces the inflammatory responses of rheumatoid arthritis synovial fibroblasts via Toll-like receptor 4. *Clin Exp Rheumatol* 2019;37:834–41.
18 Roome T, Aziz S, Razzak A, *et al*. Opuntioside, opuntiol and its metallic nanoparticles attenuate adjuvant-induced arthritis: novel suppressors of Toll-like receptors -2 and -4. *Biomed Pharmacother* 2019;112.
19 Alsousi AA, Igwe OJ. Redox-active trace metal-induced release of high mobility group box 1(HMGB1) and inflammatory cytokines in fibroblast-like synovial cells is Toll-like receptor 4 (TLR4) dependent. *Biochim Biophys Acta Mol Basis Dis* 2018;1864:3847–58.
20 Abed El-Gaphar OAM, Abo-Youssef AM, Halal GK. Levetiracetam mitigates lipopolysaccharide-induced JAK2/STAT3 and TLR4/MAPK signaling pathways activation in a rat model of adjuvant- induced arthritis. *Eur J Pharmacol* 2018;826:85–95.
21 Lee J-H, Kim B, Jin WJ, *et al*. Pathogenic roles of CXCL10 signaling through CXCR3 and TLR4 in macrophages and T cells: relevance for arthritis. *Arthritis Res Ther* 2017;19.
22 Tripathy A, Khanna S, Padhan P, *et al*. Direct recognition of LPS drive TLR4 expressing CD8+ T cell activation in patients with rheumatoid arthritis. *Sci Rep* 2017;7:933.
23 Kiyeko GW, Hatterer E, Herren S, *et al*. Spatiotemporal expression of endogenous TLR4 ligands leads to inflammation and bone erosion in mouse collagen-induced arthritis. *Eur J Immunol* 2016;46:2629–38.
24 Watanabe T, Takahashi N, Hirabara S, *et al*. Hyaluronan inhibits Tlr-4-Dependent RANKL expression in human rheumatoid arthritis synovial fibroblasts. *PLoS One* 2016;11:e0153142.
25 Blair HA, Deeks ED. Abatacept: a review in rheumatoid arthritis. *Drugs* 2017;77:1221–33.
26 Lee YH, Bae S-C. Comparative efficacy and safety of tocilizumab, rituximab, abatacept and tofacitinib in patients with active rheumatoid arthritis that inadequately responds to tumor necrosis factor inhibitors: a Bayesian network meta-analysis of randomized controlled tri. *Int J Rheum Dis* 2016;19:1103–11.
27 Kuijper TM, Lamers-Karnebeek FBG, Jacobs JWG, *et al*. Flare rate in patients with rheumatoid arthritis in low disease activity or remission when tapering or stopping synthetic or biologic DMARD: a systematic review. *J Rheumatol* 2015;42:2012–22.
28 Brown S, Everett CC, Naraghi K, *et al*. Alternative tumour necrosis factor inhibitors (TNFi) or abatacept or rituximab following failure of initial TNFi in rheumatoid arthritis: the switch RCT. *Health Technol Assess* 2018;22:1–280.
29 de Hair MJH, Jacobs JWG, Schoneveld JLM, *et al*. Difficult-To-Treat rheumatoid arthritis: an area of unmet clinical need. *Rheumatology* 2017. doi:10.1093/rheumatology/kex349. [Epub ahead of print: 04 Oct 2017].
30 Dunn-Siegrist I, Leger O, Daubeuf B, *et al*. Pivotal involvement of Fcgamma receptor IIA in the neutralization of lipopolysaccharide signaling via a potent novel anti-TLR4 monoclonal antibody 15C1. *J Biol Chem* 2007;282:34817–27.
31 Monnet E, Lapeyre G, Poelgeest Evan, *et al*. Evidence of NI-0101 pharmacological activity, an anti-TLR4 antibody, in a randomized phase I dose escalation study in healthy volunteers receiving LPS. *Clin. Pharmacol. Ther.* 2017;101:200–8.
32 Kremer JM, Cohen S, Wilkinson BE, *et al*. A phase IIb dose-ranging study of the oral JAK inhibitor tofacitinib (CP-690,550) versus placebo in combination with background methotrexate in patients with active rheumatoid arthritis and an inadequate response to methotrexate alone. *Arthritis Rheum* 2012;64:970–81.
33 Genovese MC, Bojin S, Biagini IM, *et al*. Tabalumab in rheumatoid arthritis patients with an inadequate response to methotrexate and naive to biologic therapy: a phase II, randomized, placebo-controlled trial. *Arthritis Rheum* 2013;65:880–9.
34 Cameán-Castillo M, Gimeno-Ballester V, Rios-Sanchez E, *et al*. Network meta-analysis of tofacitinib versus biologic treatments in moderate-to-severe rheumatoid arthritis patients. *J Clin Pharm Ther* 2019;44:384–96.

35 Singh JA, Hossain A, Mudano AS, *et al*. Biologics or tofacitinib for people with rheumatoid arthritis naive to methotrexate: a systematic review and network meta-analysis. *Cochrane Database Syst Rev* 2017;59.
36 Guijarro-Muñoz I, Compte M, Álvarez-Cienfuegos A, *et al*. Lipopolysaccharide activates Toll-like receptor 4 (TLR4)-mediated NF-κB signaling pathway and proinflammatory response in human pericytes. *J Biol Chem* 2014;289:2457–68.
37 Shahrara S, Park CC, Temkin V, *et al*. Rantes modulates TLR4-Induced cytokine secretion in human peripheral blood monocytes. *J Immunol* 2006;177:5077–87.
38 Elshabrawy HA, Essani AE, Szekanecz Z, *et al*. Tlrs, future potential therapeutic targets for RA. *Autoimmun Rev* 2017;16:103–13.
39 Smolen JS, Burmester G-R, Combe B, *et al*. Head-to-head comparison of certolizumab pegol versus adalimumab in rheumatoid arthritis: 2-year efficacy and safety results from the randomised EXXELERATE study. *The Lancet* 2016;388:2763–74.
40 Genovese MC, Braun DK, Erickson JS, *et al*. Safety and efficacy of open-label subcutaneous ixekizumab treatment for 48 weeks in a phase II study in biologic-naive and TNF-IR patients with rheumatoid arthritis. *J Rheumatol* 2016;43:289–97.
41 Wong YN, Rossignol D, Rose JR, *et al*. Safety, pharmacokinetics, and pharmacodynamics of E5564, a lipid A antagonist, during an ascending single-dose clinical study. *J Clin Pharmacol* 2003;43:735–42.
42 Rice TW, Wheeler AP, Bernard GR, *et al*. A randomized, double-blind, placebo-controlled trial of TAK-242 for the treatment of severe sepsis*. *Crit Care Med* 2010;38:1685–94.
43 Nicolle L. Best pharmacological practice: urinary tract infections. *Expert Opin Pharmacother* 2003;4:693–704.
44 Raz R. Urinary tract infection in postmenopausal women. *Korean J Urol* 2011;52:801–8.

CLINICAL SCIENCE

What is axial spondyloarthritis? A latent class and transition analysis in the SPACE and DESIR cohorts

Alexandre Sepriano [1,2] Sofia Ramiro [1,3] Désirée van der Heijde [1] Floris van Gaalen,[1] Pierre Hoonhout,[4] Anna Molto [5,6] Alain Saraux [7] Roberta Ramonda,[8] Maxime Dougados,[5,6] Robert Landewé [3,9]

Handling editor Josef S Smolen

► Additional material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-216516).

For numbered affiliations see end of article.

Correspondence to
Professor Robert Landewé, Amsterdam University Medical Centres, Amsterdam 1105 AZ, Netherlands;
landewe@rlandewe.nl

Received 22 October 2019
Revised 19 December 2019
Accepted 4 January 2020
Published Online First 24 January 2020

ABSTRACT

Objectives To gain expert-judgement-free insight into the *Gestalt* of axial spondyloarthritis (axSpA), by investigating its 'latent constructs' and to test how well these latent constructs fit the Assessment of SpondyloArthritis international Society (ASAS) classification criteria.

Methods Two independent cohorts of patients with early onset chronic back pain (SPondyloArthritis Caught Early (SPACE)) or inflammatory back pain (IBP) (DEvenir des Spondylarthopathies Indifférenciées Récentes (DESIR)) were analysed. Latent class analysis (LCA) was used to estimate the (unobserved) potential classes underlying axSpA. The best LCA model groups patients into clinically meaningful classes with best fit. Each class was labelled based on most prominent features. Percentage fulfilment of ASAS axSpA, peripheral SpA (pSpA) (ignoring IBP) or both classification criteria was calculated. Five-year data from DESIR were used to perform latent transition analysis (LTA) to examine if patients change classes over time.

Results SPACE (n=465) yielded four discernible classes: 'axial' with highest likelihood of abnormal imaging and HLA-B27 positivity; 'IBP+peripheral' with 100% IBP and dominant peripheral symptoms; 'at risk' with positive family history and HLA-B27 and 'no SpA' with low likelihood for each SpA feature. LCA in DESIR (n=576) yielded similar classes, except for the 'no-SpA'. The ASAS axSpA criteria captured almost all (SPACE: 98%; DESIR: 93%) 'axial' patients, but the 'IBP+peripheral' class was only captured well by combining the axSpA and pSpA criteria (SPACE: 78%; DESIR: 89%). Only 4% of 'no SpA' patients fulfilled the axSpA criteria in SPACE. LTA suggested that 5-year transitions across classes were unlikely (11%).

Conclusion The *Gestalt* of axSpA comprises three discernible entities, only appropriately captured by combining the ASAS axSpA and pSpA classification criteria. It is questionable whether some patients with 'axSpA at risk' will ever develop axSpA.

Key messages

What is already known about this subject?

- Axial spondyloarthritis (axSpA) is a disease that is difficult to diagnose; its *Gestalt* is more than only a collection of SpA features included in the Assessment of SpondyloArthritis international Society (ASAS) axSpA classification criteria.
- The ASAS axSpA criteria may suffer from inappropriate circular reasoning because they were developed against expert's opinion.

What does this study add?

- An analytical technique that circumvents expert opinion (latent class analysis) was used to determine, in a circularity-free manner, the *Gestalt* of axSpA and yielded three recognisable clinical entities labelled as: 'pure axial SpA', 'axial SpA with peripheral signs' and 'axial SpA at risk'.
- 'Pure axial SpA' represents the conventional clinical picture of axSpA and is well captured by the ASAS axSpA criteria, but patients with 'axial SpA with peripheral signs' mostly fulfil the peripheral SpA (pSpA) criteria, suggesting a larger overlap between axSpA and pSpA than anticipated when the ASAS criteria were developed.
- The 'axial SpA at risk' entity (often captured by the ASAS axSpA criteria) is a 'grey zone' entity based on the presence of risk factors for axSpA and may encompass individuals that do neither have SpA nor will ever develop it.

How might this impact on clinical practice or future developments?

- Patients with 'axial SpA with peripheral signs' are not necessarily recognised as axSpA and therefore not included in axSpA trials; 'at risk' patients, on the contrary, may be overdiagnosed and overtreated, especially if classification criteria are misused for diagnostic purposes.

INTRODUCTION

Spondyloarthritis (SpA) encompasses heterogeneous entities with common clinical, laboratory and imaging features. The full spectrum of SpA includes patients with dominant axial symptoms (axial SpA (axSpA)) and patients with dominant peripheral symptoms (peripheral SpA (pSpA)).[1] The term axSpA aggregates patients with radiographic axSpA (r-axSpA; also known as ankylosing spondylitis) and non-radiographic axSpA (nr-axSpA), differing only by the presence of radiographic sacroiliitis in the former, as defined by the modified New York (mNY) criteria.[2]

axSpA is a syndrome described by classification criteria that supposedly best reflect its inherently unmeasurable 'latent' construct (*Gestalt*). The

To cite: Sepriano A, Ramiro S, van der Heijde D, *et al*. *Ann Rheum Dis* 2020;**79**:324–331.

Assessment of SpondyloArthritis international Society (ASAS) criteria for axSpA have been developed to classify both r-axSpA and nr-axSpA. In the absence of a 'gold standard', expert opinion has been used as an external 'anchor' to develop and validate classification criteria.[3–5] The ASAS criteria outperform other criteria,[6] meaning that they contain several elements that experts consider relevant for their 'latent' picture of axSpA.

While such an approach for developing classification criteria has been pursued by default in rheumatology, it has a fundamental limitation that may jeopardise their construct and content validity: circularity. If criteria are developed against expert opinion, and the expert finds certain characteristics (eg, inflammation on MRI of the sacroiliac joints (SIJ)) more important than others, such characteristics may be awarded a too prominent place in the criteria. Subsequent cross-validation against an expert diagnosis may produce results driven by experts' beliefs rather than on an objective presence of axSpA. The *axiom* that 'early (diagnosis and treatment) is always better', a dominant view in modern rheumatology, may have contributed to rheumatologists' beliefs and as such trickled down into the ASAS criteria, designed to better capture patients with early disease. When classification criteria are (mis)used in a diagnostic context, overdiagnosis, followed by overtreatment, is a logical consequence.[7]

A more circularity-free determination of the *Gestalt* of axSpA is lacking in the literature, which hampers the study of the side effects of overdiagnosis and overtreatment. Here, we propose to evaluate the *Gestalt* of axSpA using an analytical approach that excludes the rheumatologist's diagnostic opinion. Our aims were twofold: (i) to gain an expert-judgement-free insight, into the concept of axSpA, by investigating its 'latent constructs' and (ii) to evaluate how well the ASAS SpA classification criteria capture these 'latent constructs'.

METHODS

Patients and study design

Baseline data from the SPondyloArthritis Caught Early (SPACE) cohort and baseline and 5-year data from the DEvenir des Spondylarthopathies Indifférenciées Récentes (DESIR) were used. Both cohorts have been previously described in detail.[6 8] Briefly, in SPACE (ongoing multinational cohort), consecutive patients aged ≥16 years with chronic back pain (≥3 months, ≤2 years and onset <45 years) are included. In DESIR, consecutive patients aged 18–50 years with inflammatory back pain (IBP) (>3 months but <3 years), and for whom the treating rheumatologist considers the symptoms suggestive of axSpA (level of confidence (LoC) ≥5, scale 0–10), were included. Databases were locked in October 2017 (SPACE) and June 2016 (DESIR).

SpA features

The following features were collected in each cohort: HLA-B27, elevated C reactive protein (CRP) (≥6 mg/L), family history of SpA (ASAS definition),[5] good response to non-steroidal anti-inflammatory drugs (NSAIDs), peripheral arthritis, heel enthesitis, dactylitis, psoriasis, inflammatory bowel disease, acute anterior uveitis and IBP.

At baseline, SpA features were considered positive if 'ever present' (ie, any time in the past and/or baseline) in both cohorts, except dactylitis (available only as 'current' in SPACE). In DESIR, data on SpA features were also collected every 6 months up to 2 years and yearly thereafter up to 5 years. Change in time-varying features was defined as 'once-a-feature-always-a-feature (OFAF)': patients positive at baseline remained positive at 5 years, even if becoming negative or missing in between; patients negative at baseline, remained negative at follow-up if no switch to positive or if missing in between. A feature changed to positive if appearing anytime during follow-up.

Radiographs and MRIs of the SIJ (X-SIJ; MRI-SIJ) and spine (X-Spine; MRI-Spine) were obtained at baseline in both cohorts, and at 2 and 5 years in DESIR. Each image was independently scored, by three trained central readers in each cohort, blinded to chronology, clinical data and to the results of other modalities. Four binary imaging features, defined by agreement between ≥2 out of 3 readers, were assessed: inflammation on MRI-SIJ (ASAS definition)[9 10]; bone marrow oedema (BME) on MRI-Spine (≥5 lesions)[11]; definite structural damage in X-SIJ according to the mNY criteria[2] and ≥1 syndesmophyte in X-Spine.[12]

Statistical analysis

Latent class analysis (LCA) was performed with baseline data of each cohort separately, including patients with complete data on all features. LCA unmasks a 'latent' (ie, unobserved) construct (here: *Gestalt* of axSpA) by splitting patients into mutually exclusive classes based on the covariance of observed SpA features. Extensive evidence supports the superiority of LCA in identifying latent data structures, compared with other clustering methods.[13–15] SpA features (15 variables in SPACE; 14 in DESIR (excluding IBP)) were selected 'a priori' based on content knowledge without predefined weights.

A detailed description LCA and how it can be used to identify the latent classes of the *Gestalt* of axSpA is provided in online supplementary text S1. Briefly, the number of classes was increased, one-by-one, until the best model was found, defined by: best goodness of fit assessed by Akaike's information criterion, Bayesian information criterion (BIC), sample-sized adjusted BIC, entropy, likelihood ratio test (comparing the model with the one with n-1 classes) and by clinically recognisable patterns within each class (ie, a statistical criterion alone does not suffice). The classes of the final model were interpreted according to the probability of each feature and labelled as a clinically recognisable entity. Features were defined as: across-class dominant (highest probability across classes); within-class dominant (probability >50% within each class) and not dominant across or within classes.

Maximum likelihood estimates were used to classify individual patients based on their posterior probability of class membership. This allowed us to describe the classes including also variables not used in the models and to evaluate the percentage of patients within each class fulfilling the ASAS axSpA, pSpA (ignoring IBP) and the SpA criteria (ie, combination of either axSpA or pSpA criteria) at baseline.

To address between-cohort differences in study design, a sensitivity analysis was performed in SPACE: only in patients with a rheumatologist's diagnosis with LoC ≥5 (similar to DESIR).

Latent transition analysis (LTA) was used to estimate the likelihood of change across classes after 5 years in DESIR.[16] LTA includes the same patients and variables as in LCA. The number of classes best fitting the baseline and 5-year LCA formed the basis of the LTA model. Classes at baseline and follow-up can be assumed as: having the same meaning (full invariance); different meaning (full non-invariance) or the same meaning for some and different for others (partial invariance). The final LTA model has the number of classes at baseline and 5 year and class-(in)variance that best fits the data provided it is clinically meaningful.

LCA was performed in Stata V.15.1. LTA was performed in MPlus V.7.

Table 1 Baseline patient characteristics in the SPACE and DESIR cohorts

	SPACE (n=465)	DESIR (n=576)
Age at baseline (years)	30 (8)	33 (8)
Male gender	126 (45)	269 (47)
Symptom duration (years)	1.8 (2.0)	1.5 (0.8)
ASAS axSpA criteria	172 (37)	358 (62)
axSpA according to rheumatologist*	136 (52)	269 (47)
ASAS pSpA criteria	182 (39)	320 (56)
ASAS SpA criteria†	249 (54)	443 (77)
Sacroiliitis on MRI-SIJ (ASAS)	64 (14)	153 (27)
BME on MRI-Spine (≥5 lesions)	21 (5)	25 (4)
Radiographic sacroiliitis (mNY)	38 (8)	78 (14)
≥1 syndesmophyte on X-Spine	15 (3)	39 (7)
Elevated CRP (≥6 mg/L)	64 (24)	169 (29)
Good response to NSAIDs ever	189 (41)	491 (85)
Peripheral arthritis ever	41 (15)	122 (21)
Dactylitis ever	18 (6)	78 (14)
Heel enthesitis ever	55 (20)	261 (45)
HLA-B27	156 (57)	345 (60)
Family history of SpA	123 (44)	250 (43)
Psoriasis ever	41 (15)	99 (17)
Uveitis ever	27 (10)	52 (9)
Inflammatory bowel disease ever	16 (6)	25 (4)
Current arthritis/any enthesitis/dactylitis	317 (68)	398 (69)
Inflammatory back pain	308 (66)	576 (100)
Number of SpA features (0–9)‡	2 (1)	3 (1)

Values are mean (SD) for continuous variables or number (%) for binary variables. SpA features are positive if 'ever present' (any time in the past and/or baseline).
*Clinical diagnosis of axSpA at baseline with a level of confidence >7; missing data SPACE: symptom duration (n=461); missing data DESIR: axSpA according to rheumatologist (n=575).
†Fulfilment of either ASAS axSpA or ASAS pSpA classification criteria.
‡Peripheral arthritis, heel enthesitis, dactylitis, psoriasis, uveitis, inflammatory bowel disease, good response to NSAIDs, elevated CRP and family history of SpA.
ASAS, Assessment of SpondyloArthritis international Society; axSpA, axial spondyloarthritis; BME, bone marrow oedema; CRP, C reactive protein; mNY, modified New York criteria; NSAIDs, non-steroidal anti-inflammatory drugs; pSpA, peripheral spondyloarthritis; SIJ, sacroiliac joints; X-Spine, radiograph of the spine.

RESULTS

Baseline characteristics

In total, 465 patients from SPACE and 576 from DESIR were included. In SPACE, included patients were more likely to be HLA-B27 positive (57% vs 37%) and less likely to have BME on MRI-SIJ (14% vs 30%) than those excluded (n=283). No differences were seen in DESIR (excluded: n=132) (online supplementary tables S1 and S2). Baseline characteristics of the included patients from both cohorts are shown in table 1. Patients from DESIR had, on average, more SpA features compared with those from SPACE, including peripheral features (eg, heel enthesitis 45% vs 20%) and axial imaging abnormalities (eg, sacroiliitis on MRI-SIJ 27% vs 14%).

Latent class analysis in SPACE and DESIR

A 4-class (SPACE) and a 3-class (DESIR) LCA model fitted the data best (table 2). The additional class in the 5-class (SPACE) and 4-class (DESIR) models, with worse model fit, did not yield a clinically recognisable pattern (online supplementary tables S3, S4 and S5).

The final LCA models are shown in table 2. In SPACE, class 1 was characterised by highest likelihood (ie, across-class dominance) of lesions present on axial imaging, elevation of CRP and HLA-B27 positivity, and was labelled as 'axial'. Class 2 was labelled 'IBP+peripheral', given the 100% likelihood of IBP and across-class dominance of peripheral features. Class 3 had across-class dominance of positive family history (71%) and within-class dominance of HLA-B27 positivity (69%) and IBP (66%) but low likelihood of other features and was labelled as 'at risk'. Class 4 was labelled 'no SpA' given the very low likelihood for each SpA feature.

The LCA analysis in DESIR yielded the same latent classes, except 'no SpA', and an overlapping pattern of dominance: among 42 possible comparisons (14 features (excluding IBP) multiplied by 3 classes (excluding 'no SpA')), in 37 (88%) the dominance pattern was similar to SPACE (table 2). Figure 1 graphically displays the between-cohort similarities, and also the phenotypical differences between the 'axial' and 'IBP+peripheral' classes which overlap with the 'at risk class' only partially, and even less with the 'no SpA' class.

The LCA model in SPACE, in patients with a rheumatologist's diagnosis of axSpA (LoC ≥5) (n=202) yielded the same classes as the main model, except 'no SpA' that is similar to DESIR ('axial': 29%; 'IBP+peripheral': 33%; 'at risk':38%; online supplementary table S6).

Latent transition analysis in DESIR

Of the 576 patients in DESIR, 500 (87%) completed the 5-year follow-up. The change in SpA and imaging features between baseline and 5 years is shown in figure 2A. Because of how SpA features were defined (OFAF), all increased in prevalence over time, but changes were more pronounced with peripheral (eg, peripheral arthritis: 21%–30%) than with imaging features (eg, BME on MRI-SIJ: 26%–29%).

Similar to baseline LCA, a 3-class model at 5 years best fitted the data (online supplementary table S7 and S8). Accordingly, an LTA model with three classes at both timepoints was fit. Although the model fit (online supplementary table S9) was better with partial invariance, the resulting model did not yield a clinically recognisable pattern (data not shown), so the simplest assumption (full invariance) was taken to define the final LTA model (figure 2b, online supplementary table S10). LTA revealed a 0% probability of switch from the 'axial' and 'IBP+peripheral' to another class. 'At risk' patients at baseline had 11% likelihood to change to 'IBP+peripheral' over 5 years.

Observed characteristics and fulfilment of the ASAS classification criteria

The patterns of observed characteristics per latent class in SPACE and DESIR were, expectedly, similar to the model-based estimates (table 3). In addition, across-class dominance of males in the 'axial' class (SPACE: 66%; DESIR: 73%), and current arthritis/enthesitis/dactylitis (ie, entry criterion for pSpA criteria) in the 'IBP+peripheral class (SPACE: 87%; DESIR: 88%) were observed.

The ASAS axSpA criteria captured almost all patients from the 'axial' class in SPACE (63/64; 98%). This percentage was much lower with 'IBP+peripheral' (41/92; 49%), and missed patients were most often female (78%), positive for current arthritis/enthesitis/dactylitis (92%) and HLA-B27 and MRI-SIJ/mNY negative. The pSpA criteria captured 67% of the 'IBP+peripheral' patients and this figure was 78% when the axSpA and pSpA criteria were combined. Fifty-nine (60%) patients from the

Table 2 Final latent class analysis (LCA) models in SPACE (n=465) and DESIR (n=576) in probability scale (range: 0–1)

	SPACE				DESIR			
	Class 1 ('axial') (p*=16%)	Class 2 ('IBP+peripheral') (p*=20%)	Class 3 ('at risk') (p*=24%)	Class 4 ('no SpA') (p*=40%)	Class 1 ('axial') (p*=19%)	Class 2 ('IBP+peripheral') (p*=27%)	Class 3 ('at risk') (p*=54%)	Class 4 'no SpA'†
Inflammation on MRI-SIJ (ASAS)	0.74	0.04	0.00	0.03	0.83	0.22	0.09	
BME on MRI-Spine (≥5 lesions)	0.25	0.02	0.00	0.00	0.20	0.00	0.01	
Radiographic sacroiliitis (mNY)	0.32	0.09	0.01	0.03	0.58	0.06	0.02	
≥1 syndesmophyte on X-Spine	0.03	0.06	0.00	0.04	0.11	0.05	0.06	
Elevated CRP (≥6 mg/dL)	0.49	0.22	0.21	0.20	0.56	0.41	0.14	
Good response to NSAIDs (ever)	0.59	0.85	0.25	0.20	0.97	0.84	0.82	
Peripheral arthritis (ever)	0.17	0.44	0.04	0.10	0.09	0.73	0.00	
Dactylitis (ever)	0.02	0.18	0.00	0.03	0.03	0.46	0.01	
Heel enthesitis (ever)	0.10	0.66	0.13	0.04	0.26	0.60	0.45	
HLA-B27	0.84	0.33	0.69	0.00	0.90	0.52	0.53	
Family history of SpA	0.38	0.50	0.71	0.21	0.48	0.44	0.41	
Psoriasis (ever)	0.10	0.31	0.02	0.08	0.09	0.29	0.14	
Uveitis (ever)	0.13	0.07	0.12	0.02	0.08	0.12	0.08	
IBD (ever)	0.03	0.15	0.00	0.10	0.02	0.05	0.05	
Inflammatory back pain	0.68	1.00	0.66	0.49	NA	NA	NA	

The table displays the main results of the LCA separately in each cohort. Values are the conditional probability for each SpA feature positivity within each latent class (range: 0–1).

Heatmap legend: *red:* highlights dominant features across latent classes; *brown:* highlights dominant features (probability >50%) within each class but not across classes; *blank:* not dominant neither across nor within classes. SpA features are positive if 'ever present' (any time in the past and/or baseline).

*Probability of the latent class.

†'No SpA' latent class absent in DESIR; in DESIR all included patients have a high likelihood of axSpA.

ASAS, Assessment of SpondyloArthritis international Society; BME, bone marrow oedema; CRP, C reactive protein; DESIR, DEvenir des Spondylarthopathies Indifférenciées Récentes; IBD, inflammatory bowel disease; mNY, modified New York criteria; NA, not applicable; NSAIDs, non-steroidal anti-inflammatory drugs; SIJ, sacroiliac joints; X-Spine, radiograph of the spine.

'at risk' class fulfilled the axSpA criteria (58/59=98% fulfilling the 'clinical arm only'). Among the 58 fulfilling the 'clinical arm only', family history of SpA (75%) and IBP (85%) were the most common features. Only nine patients (4%) from the 'no SpA' class fulfilled the axSpA criteria, all of which captured by the imaging arm only (78% positive for IBP or good response to NSAIDs). Results were similar in DESIR, except that the percentage of 'at Risk' patients fulfilling the 'clinical arm only' was somewhat lower (148/177=84%).

DISCUSSION

Using a data-driven approach, we identified three separate clinical entities, remarkably stable over time, together forming the

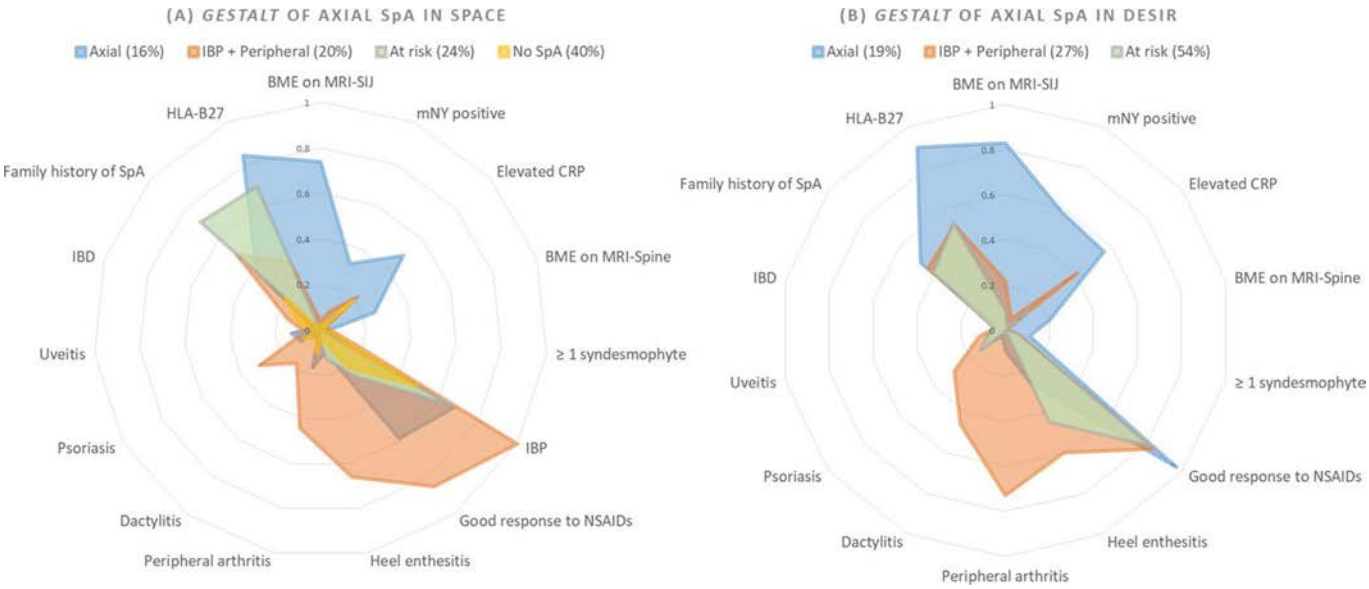


Figure 1 Radar charts showing the distribution of the probabilities of each feature according to the final LCA model in (A) SPACE and (B) DESIR. BME, bone marrow edema; CRP, C reactive protein; DESIR, DEvenirdes Spondylarthopathies Indifférenciées Récentes; IBD, inflammatory bowel disease; IBP, inflammatory back pain; mNY, modified New York criteria; NSAIDs, non-steroidal anti-inflammatory drugs; SIJ, sacroiliac joints; SPACE, SPondyloArthritisCaught Early; MRI-SIJ, magnetic ressonance imaging of the sacroiliac joints; MRI-Spine, MRI of the spine.

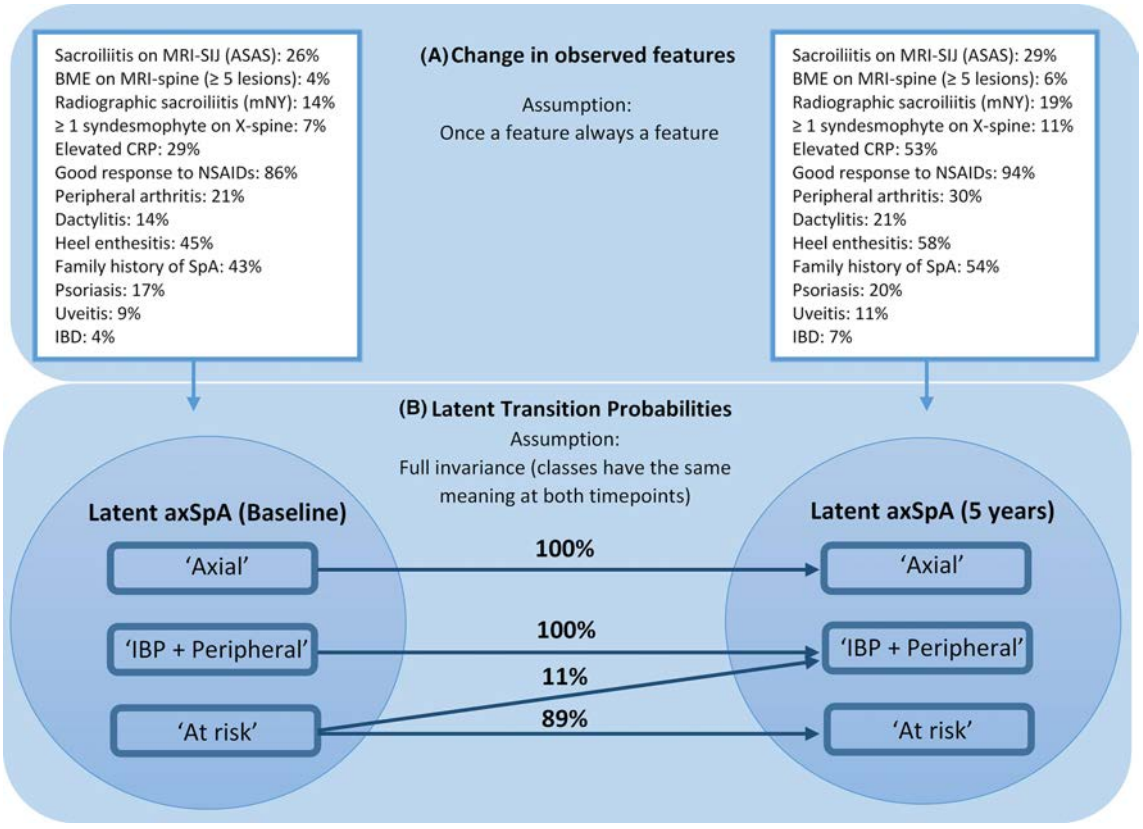


Figure 2 Final latent transition analysis (LTA) model (with full invariance*) in DEvenirdes Spondylarthopathies Indifférenciées Récentes (n=576). (A) Squares refer to observed (ie, measurable) variables and (B) Circles refer to latent (ie, unobserved) variables. Arrows: latent transition analysis models the change in observed features (A) to estimate the latent (B) transition probabilities between classes from baseline to 5 years. ASAS, Assessment of SpondyloArthritis international Society; axSpA, axial spondyloarthritis; BME, bone marrow edemao; CRP, C reactive protein; IBD, inflammatory bowel disease; IBP, inflammatory back pain; LTA, latent transition analysis; mNY, modified New York criteria; NSAIDs, non-steroidal anti-inflammatory drug; SIJ, sacroiliac joints; X-Spine, radiograph of the spine MRI, magnetic ressonance imaging. *Selection of final LTA model according to goodness of fit detailed in online supplementary table S9 and full final model in online supplementary table S10.

Gestalt of axSpA, in two independent cohorts, that we labelled 'pure axial SpA' ('axial'), 'axial SpA with peripheral signs' ('IBP+peripheral') and 'axial SpA at risk' ('at risk'). In SPACE, a cohort that includes patients with back pain without axSpA, these three axSpA classes decently discerned themselves from a fourth labelled as 'no SpA'. This adds to the credibility of our data, since the absence of 'no SpA' in DESIR was expected based on enrolment criteria. The ASAS axSpA classification criteria captured almost entirely the 'axial' class but missed several patients from the 'IBP+peripheral' class: the latter is better captured when combining the axSpA and pSpA criteria, suggesting a larger overlap between axSpA and pSpA than previously thought, when the ASAS criteria were developed. Taken together, at the group level these results confirm the robustness of the classification criteria. The 'at risk' class is an entity characterised by the presence of presumed risk factors for axSpA but the absence of objective clinical signs. While these patients often fulfil the ASAS axSpA classification criteria, it is likely that some do not actually have or will ever develop axSpA. Overdiagnosis of axSpA in the 50% of patients in this class is likely if classification criteria are ticked for diagnosis.

A diagnosis of axSpA is challenging and should rely on thorough knowledge and recognition of 'the appropriate pattern'.[17 18] The rheumatologists' perception of the 'SpA pattern' evolved over the last 40 years as a result of efforts by the international rheumatology community. Initially, only r-axSpA (ankylosing spondylitis) was recognised and classified by the mNY criteria.[2] In the 70s–80s, Moll and Wright defined SpA as a group of entities with common features,[19] and the Amor and the European Spondyloarthropathy Study Group (ESSG) classification criteria were proposed.[20 21] Both criteria sets capture the broader 'SpA pattern' by combining axial and peripheral features and do not distinguish between patients with dominant axial and dominant peripheral patterns. Since then, evidence has emerged supporting that patients with the axial and peripheral pattern may respond differently to treatment,[22 23] and that not all patients with axSpA will develop sacroiliitis on pelvic radiographs (mNY-positive). When they do, this is frequently a late and unreliable finding and often preceded by sacroiliitis on MRI-SIJ for many years.[24–31] Such evidence prompted ASAS experts to develop classification criteria for patients with predominant axial involvement,[5] also capturing those that are mNY-negative (nr-axSpA) as axSpA, and for patients with predominant peripheral involvement that—if combined—enclose the entire *Gestalt* of SpA according to experts.[4]

The ASAS axSpA and pSpA classification criteria were validated against an external 'gold standard': expert opinion.[3–5] Extensive evidence supports that the ASAS criteria perform well against this anchor,[32] but misclassification remains a matter of intense debate.[33] It has been argued that expert opinion may have contributed to designing criteria that encompass circular reasoning,[34 35] that is, features deemed important by experts, especially those that allow early detection (eg, sacroiliitis on MRI), were awarded a too prominent place in criteria that were

Table 3 Baseline observed patient and disease characteristics per latent class in SPACE and DESIR

	SPACE				DESIR		
	'Axial' (n=64)	'IBP+peripheral' (n=92)	'At risk' (n=99)	'No SpA' (n=210)	'Axial' (n=110)	'IBP+peripheral' (n=137)	'At risk' (n=329)
Clinical and demographic							
Age at baseline (years)	30	32	30	31	31	33	34
Male gender	***66***	38	32	25	***73***	43	40
Symptom duration (years)	1	2	2	2	2	1	2
Imaging							
Inflammation on MRI-SIJ (ASAS)	***86***	3	0	2	***88***	22	8
BME on MRI-Spine (≥5 lesions)	**28**	2	0	1	**20**	0	1
Radiographic sacroiliitis (mNY)	**34**	11	1	2	***59***	5	2
≥1 syndesmophyte on X-Spine	3	7	0	3	11	6	6
SpA features							
Elevated CRP (≥6 mg/dL)	**50**	22	24	20	***56***	39	16
Good response to NSAIDs (ever)	*59*	***89***	32	18	***97***	*84*	82
Peripheral arthritis (ever)	17	**47**	3	9	7	***83***	0
Dactylitis (ever)	2	**19**	0	2	3	***55***	0
Heel enthesitis (ever)	9	***72***	12	3	24	***59***	47
Current arthritis/enthesitis/dactylitis	48	***87***	*64*	*68*	*56*	***88***	*65*
HLA-B27	***86***	36	***85***	0	***93***	*52*	*52*
Family history of SpA	38	*52*	***72***	24	47	43	42
Psoriasis (ever)	9	**35**	1	7	7	**29**	16
Uveitis (ever)	13	7	**15**	2	8	**12**	8
IBD (ever)	3	**15**	0	9	1	5	5
Inflammatory back pain	*67*	***100***	*69*	50	100*	100*	100*
Number of SpA features (0–9)†	3	**5**	2	1	3	**4**	2
ASAS classification criteria							
ASAS axSpA criteria	**98**	45	60	4	**93**	58	54
ASAS pSpA criteria	48	**70**	56	15	56	**82**	44
ASAS SpA criteria‡	**98**	78	79	17	**99**	89	64

Values are means for continuous variables or percentages for binary variables.
*By study design all patients in DESIR have IBP.
†Peripheral arthritis, heel enthesitis, dactylitis, psoriasis, uveitis, inflammatory bowel disease, good response to NSAIDs, elevated CRP and family history of SpA.
‡Fulfilment of either ASAS axSpA or ASAS pSpA classification criteria. Values in bold highlight discriminant features across latent classes. Values in italic highlight dominant features (probability >50%) within each class. SpA features are positive if 'ever present' (any time in the past and/or baseline).
ASAS, Assessment of SpondyloArthritis international Society; axSpA, axial spondyloarthritis; BME, bone marrow oedema; CRP, C reactive protein; DESIR, DEvenir des Spondylarthopathies Indifférenciées Récentes; IBD, inflammatory bowel disease; mNY, modified New York criteria; NSAIDs, non-steroidal anti-inflammatory drug; pSpA, peripheral spondyloarthritis; SIJ, sacroiliac joints; SPACE, SPondyloArthritis Caught Early; X-Spine, radiograph of the spine.

subsequently again validated by experts. However, whether or not circularity has played a decisive role remains unclear, since an expert-judgement-free assessment of the *Gestalt* of axSpA has not been pursued so far. This is exactly what we have done in this study.

Using LCA we could describe the *Gestalt* of axSpA without any pre-assumptions on the contribution ('weight') of each SpA feature. This was only possible because LCA, following selection of parameters for analysis, does not need interpretational input from experts, whose beliefs therefore do not influence the analysis. The only inevitable influence experts potentially had was deciding if the patient should be included in the cohort. One of the phenotypes that arose from this analytical framework was a syndrome characterised by a high likelihood of axial imaging abnormalities, HLA-B27 positivity and male dominance, which we have subsequently labelled as 'axial'. This phenotype closely resembles the rheumatologist's conventional clinical picture of axSpA. Of note, LCA did not distinguish nr-axSpA from r-axSpA, even after forcing one additional class to the model. This is in line with extensive evidence suggesting that the split of axSpA in nr-axSpA and r-axSpA is artificial and supports the view that both are part of the same disease spectrum.[1 26 36 37]

However, the 'axial' class is only one part of the *Gestalt* of axSpA: we identified a separate phenotype, defined by the presence of IBP (100%) in close conjunction with peripheral signs and symptoms ('IBP+peripheral'). These patients with axSpA (mostly female) had back pain but were unlikely to be positive for sacroiliitis on imaging and HLA-B27. Thus, these patients rather fulfilled the pSpA than the axSpA classification criteria since the latter require either positive imaging ('imaging arm') or HLA-B27 ('clinical arm'). Formally, the ASAS pSpA criteria could not have been applied, since all patients had IBP.[4] We ignored this rule to better understand the possible overlap between *SpA* with predominantly peripheral features (original 'target' of the pSpA criteria) and *axSpA* with peripheral signs (the entity described here). The high percentage of 'IBP+peripheral' patients fulfilling the pSpA criteria argues in favour of a significant overlap. This is in line with another study in DESIR, in which a different analytical approach (cluster analysis) was pursued that, unlike LCA, assumes an a priori presence of subgroups.[38] Taken all together, our findings undermine the current stand that either sacroiliitis on imaging or presence of HLA-B27 is mandatory to classify patients as axSpA. Several (female) patients presenting with IBP and concomitant peripheral manifestations but without manifest

sacroiliitis or HLA-B27 are not recognised as axSpA and therefore not included in axSpA trials. These patients have consistently shown to have significant burden of disease.[38–41] Whether or not these patients truly have inflammatory SpA or rather a chronic pain syndrome is a question that cannot be resolved by this analysis.

A third phenotype we identified is based on the presence of risk factors for axSpA (ie, positive family history and HLA-B27) in association with IBP and only sporadically other SpA features. We have labelled this phenotype axSpA 'at risk'. Here, 'at risk' means that patients present with features suggestive of axSpA, but such a diagnosis is not beyond any doubt. In other words, the 'at risk' class implies a higher level of uncertainty (grey zone) than the other classes, such as the 'axial' and the 'IBP+peripheral' classes. Too often, when dealing with uncertain or difficult cases clinicians apply classification criteria to inform binary diagnostic judgements (eg, axSpA vs no axSpA) that do not allow grey zones. In addition, the anchoring features of this class (ie, family history and HLA-B27) have shown redundancy,[42] but yet count separately for classification, which may contribute to overcalling axSpA when the ASAS axSpA criteria are wrongly used for diagnostic purposes. The high likelihood of IBP in these patients does not further help in discriminating SpA and no-SpA, since it also occurs in half of the patients of the 'no SpA' class. This is in line with recent data suggesting that specificity of IBP is lower than previously thought.[43 44] Although a longer follow-up may reveal more across-class switches over time, the low likelihood of 'at risk' patients to switch to a more profound phenotype within 5 years adds to the notion that 'at risk' patients may not have 'real' axSpA and will most often also not develop it later. A logical consequence would be to refrain from treating them as if they really have axSpA and from including these 'at risk' patients in axSpA trials which is indeed done as in addition to fulfilment of the ASAS criteria objective signs of inflammation are required.

In summary, we identified three latent phenotypes of the *Gestalt* of axSpA with a method that largely circumvents the circularity by expert opinion. 'Pure axial SpA' is the 'classical' phenotype of axSpA. 'axSpA with peripheral signs' is a recognisable phenotype in the spectrum of patients presenting with chronic back pain, best captured by the pSpA criteria suggesting that the overlap between axSpA and pSpA is larger than anticipated. The 'at risk' class is the least well-defined of all entities and may encompass individuals at risk of axSpA, but without fully established disease, and also individuals who do not have SpA or will ever develop it. Studies addressing the prognosis of these subphenotypes, especially that of the 'at risk' class, should inform us better on the real outcome of axSpA.

Author affiliations
[1]Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
[2]NOVA Medical School, Universidade Nova de Lisboa, Lisbon, Portugal
[3]Zuyderland Medical Centre Heerlen, Heerlen, The Netherlands
[4]Lisbon School of Economics and Management, Lisbon, Portugal
[5]Department of Rheumatology, Hôpital Cochin. Assistance Publique - Hôpitaux de Paris, Paris, France
[6]Clinical Epidemiology and Biostatistics, INSERM U1153, Paris, France
[7]Rheumatology Unit and UMR1227, Lymphocytes B et Autoimmunit, Université de Brest, Brest, France
[8]Rheumatology Unit, Department of Medicine DIMED, University of Padova, Padova, Italy
[9]Amsterdam University Medical Center, Amsterdam, The Netherlands

Twitter Alexandre Sepriano @AlexSepriano, Sofia Ramiro @sofiaramiro82 and Anna Molto @annamolto

Acknowledgements The authors would like to thank all patients who participate(d) in the SPACE cohort, and the physicians who contributed to the data collection. The authors would like to thank the different participating centres: Pr Désirée van der Heijde and Dr Floris van Gaalen (Leiden University Medical Center, Leiden, The Netherlands); Pr Robert Landewé and Dr Marleen van der Sande (Amsterdam University Medical Center, Amsterdam, The Netherlands); Dr Maikel van Oosterhout (Groene Hart Ziekenhuis Hospital, Gouda, The Netherlands); Dr Inger Jorid Berg (Diakonhjemmet Hospital, Oslo, Norway); Pr Roberta Ramonda (University of Padova, Padova, Italy); Pr Lennart Jacobsson (University of Gothenburg, Gothenburg, Sweden). The authors would also like to thank the different regional participating centres in DESIR: Pr Maxime Dougados (Paris-Cochin B), Pr Andre Kahan (Paris-Cochin A), Pr Philippe Dieudé (Paris-Bichat), Pr Bruno Fautrel (Paris-La Pitie-Salpetriere), Pr Francis Berenbaum (Paris-Saint-Antoine), Pr Pascal Claudepierre (Creteil), Pr Maxime Breban (Boulogne-Billancourt), Dr Bernadette Saint-Marcoux (Aulnay-sous-Bois), Pr Philippe Goupille (Tours), Pr Jean Francis Maillefert (Dijon), Dr Emmanuelle Dernis (Le Mans), Pr Daniel Wendling (Besancon), Pr Bernard Combe (Montpellier), Pr Liana Euller-Ziegler (Nice), Pr Pascal Richette (Paris Lariboisiere), Pr Pierre Lafforgue (Marseille), Dr Patrick Boumier (Amiens), Pr Martin Soubrier (Clermont-Ferrand), Dr Nadia Mehsen (Bordeaux), Pr Damien Loeuille (Nancy), Pr Rene-Marc Flipo (Lille), Pr Alain Saraux (Brest), Pr Xavier Mariette (LeKremlin-Bicetre), Pr Alain Cantagrel (Toulouse), Pr Olivier Vittecoq (Rouen). The authors would also like to thank the research nurses, the staff members of the Clinical Research Unit of Paris Centre, the staff members of the Biological Resource Center of Bichat Hospital, the staff members of the Department of Statistics of Nimes and all the investigators and in particular Jerome Allain, Emmanuelle Dernis, Salah Ferkal, Clement Prati, Marie-Agnes Timsit, Eric Toussirot for active patient recruitment and monitoring.

Contributors AS, SR, DvdH and RBML designed the study. AS performed the statistical analyses. AS wrote the first draft of the manuscript. All authors critically interpreted the results, reviewed the draft version and approved the final manuscript.

Funding The DESIR study is conducted as a Programme Hospitalier de Recherche Clinique with Assistance Publique Hopitaux de Paris as the sponsor. The DESIR study is also under the umbrella of the French Society of Rheumatology, which financially supports the cohort. An unrestricted grant from Pfizer has been allocated for the first 10 years. The DESIR cohort is conducted under the control of Assistance publique Hopitaux de Paris via the Clinical Research Unit Paris Centre and under the umbrella of the French Society of Rheumatology and Institut national de la sante et de la recherche medicale (Inserm). Database management is performed within the Department of Epidemiology and Biostatistics (Professeur Jean-Pierre Daures, D.I.M., Nimes, France). AS is supported by a doctoral grant from 'Fundação para a Ciência e Tecnologia' (SFRH/BD/108246/2015).

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The study was conducted according to Good Clinical Practice guidelines and approved by the appropriate local ethical committees. Written informed consent was obtained from participating patients before inclusion.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iDs
Alexandre Sepriano http://orcid.org/0000-0003-1954-0229
Sofia Ramiro http://orcid.org/0000-0002-8899-9087
Désirée van der Heijde http://orcid.org/0000-0002-5781-158X
Anna Molto http://orcid.org/0000-0003-2246-1986
Alain Saraux http://orcid.org/0000-0002-8454-7067
Robert Landewé https://orcid.org/0000-0002-0577-6620

REFERENCES

1 Sieper J, van der Heijde D. Review: Nonradiographic axial spondyloarthritis: new definition of an old disease? *Arthritis & Rheumatism* 2013;65:543–51.
2 van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the new York criteria. *Arthritis Rheum* 1984;27:361–8.
3 Rudwaleit M, Landewe R, van der Heijde D, *et al*. The development of assessment of spondyloarthritis International Society classification criteria for axial spondyloarthritis (Part I): classification of paper patients by expert opinion including uncertainty appraisal. *Ann Rheum Dis* 2009;68:770–6.
4 Rudwaleit M, van der Heijde D, Landewe R, *et al*. The assessment of spondyloarthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. *Ann Rheum Dis* 2011;70:25–31.
5 Rudwaleit M, van der Heijde D, Landewe R, *et al*. The development of assessment of spondyloarthritis International Society classification criteria for axial spondyloarthritis (Part II): validation and final selection. *Ann Rheum Dis* 2009;68:777–83.

6 van den Berg R, de Hooge M, van Gaalen F, *et al*. Percentage of patients with spondyloarthritis in patients referred because of chronic back pain and performance of classification criteria: experience from the spondyloarthritis caught early (space) cohort: table 3. *Rheumatology* 2015;54:1336.
7 Landewe RBM. Overdiagnosis and overtreatment in rheumatology: a little caution is in order. *Ann Rheum Dis* 2018;77:1394–6.
8 Dougados M, d'Agostino M-A, Benessiano J, *et al*. The DESIR cohort: a 10-year follow-up of early inflammatory back pain in France: study design and baseline characteristics of the 708 recruited patients. *Joint Bone Spine* 2011;78:598–603.
9 Lambert RGW, Bakker PAC, van der Heijde D, *et al*. Defining active sacroiliitis on MRI for classification of axial spondyloarthritis: update by the ASAS MRI Working group. *Ann Rheum Dis* 2016;75:1958–63.
10 Rudwaleit M, Jurik AG, Hermann K-GA, *et al*. Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. *Ann Rheum Dis* 2009;68:1520–7.
11 de Hooge M, van den Berg R, Navarro-Compán V, *et al*. Patients with chronic back pain of short duration from the space cohort: which MRI structural lesions in the sacroiliac joints and inflammatory and structural lesions in the spine are most specific for axial spondyloarthritis? *Ann Rheum Dis* 2016;75:1308–14.
12 Creemers MCW*et al*. Assessment of outcome in ankylosing spondylitis: an extended radiographic scoring system. *Ann Rheum Dis* 2005;64:127–9.
13 Brusco MJ, Shireman E, Steinley D. A comparison of latent class, k-means, and K-median methods for clustering dichotomous data. *Psychol Methods* 2017;22:563–80.
14 Latent class models for clustering : a comparison with K-means 2002.
15 Schreiber JB. Latent class analysis: an example for reporting results. *Research in Social and Administrative Pharmacy* 2017;13:1196–201.
16 Lanza ST, Patrick ME, Maggs JL. Latent transition analysis: benefits of a latent variable approach to modeling transitions in substance use. *J Drug Issues* 2010;40:93–120.
17 Rudwaleit M, Feldtkeller E, Sieper J. Easy assessment of axial spondyloarthritis (early ankylosing spondylitis) at the bedside. *Ann Rheum Dis* 2006;65:1251–2.
18 Rudwaleit M*et al*. How to diagnose axial spondyloarthritis early. *Ann Rheum Dis* 2004;63:535–43.
19 Moll JMH, Haslock IAN, Macrae IF, *et al*. Associations between ankylosing spondylitis, psoriatic arthritis, REITER'S disease, the intestinal arthropathies, and BEHCET'S syndrome. *Medicine* 1974;53:343–64.
20 Amor B, Dougados M, Mijiyawa M. Criteria of the classification of spondylarthropathies]. *Rev Rhum Mal Osteoartic* 1990;57:85–9.
21 Dougados M, Linden SVD, Juhlin R, *et al*. The European Spondylarthropathy Study Group preliminary criteria for the classification of Spondylarthropathy. *Arthritis & Rheumatism* 1991;34:1218–27.
22 Zochling J*et al*. Current evidence for the management of ankylosing spondylitis: a systematic literature review for the ASAS/EULAR management recommendations in ankylosing spondylitis. *Ann Rheum Dis* 2006;65:423–32.
23 Gossec L*et al*. Efficacy of cyclo-oxygenase-2 inhibition by etoricoxib and naproxen on the axial manifestations of ankylosing spondylitis in the presence of peripheral arthritis. *Ann Rheum Dis* 2005;64:1563–7.
24 Rudwaleit M, Khan MA, Sieper J. The challenge of diagnosis and classification in early ankylosing spondylitis: do we need new criteria? *Arthritis Rheum* 2005;52:1000–8.
25 Sepriano A, Rudwaleit M, Sieper J, *et al*. Five-Year follow-up of radiographic sacroiliitis: progression as well as improvement? *Ann Rheum Dis* 2016;75:1262–3.
26 van den Berg R, Lenczner G, Feydy A, *et al*. Agreement between clinical practice and trained central reading in reading of Sacroiliac joints on plain pelvic radiographs: results from the DESIR cohort. *Arthritis & Rheumatology* 2014;66:2403–11.
27 van Tubergen A*et al*. Radiographic assessment of sacroiliitis by radiologists and rheumatologists: does training improve quality? *Ann Rheum Dis* 2003;62:519–25.
28 Weber U, Lambert RGW, Østergaard M, *et al*. The diagnostic utility of magnetic resonance imaging in spondylarthritis: an international multicenter evaluation of one hundred eighty-seven subjects. *Arthritis & Rheumatism* 2010;62:3048–58.
29 Dougados M, Demattei C, van den Berg R, *et al*. Rate and predisposing factors for Sacroiliac joint radiographic progression after a two-year follow-up period in recent-onset spondyloarthritis. *Arthritis & Rheumatology* 2016;68:1904–13.
30 Dougados M, Sepriano A, Molto A, *et al*. Sacroiliac radiographic progression in recent onset axial spondyloarthritis: the 5-year data of the DESIR cohort. *Ann Rheum Dis* 2017;76:1823–8.
31 Sepriano A, Ramiro S, Landewe R, *et al*. Is active sacroiliitis on MRI associated with radiographic damage in axial spondyloarthritis? real-life data from the ASAS and DESIR cohorts. *Rheumatology* 2018.
32 Sepriano A, Rubio R, Ramiro S, *et al*. Performance of the ASAS classification criteria for axial and peripheral spondyloarthritis: a systematic literature review and meta-analysis. *Ann Rheum Dis* 2017;76:886–90.
33 Dubreuil M, Deodhar AA. Axial spondyloarthritis classification criteria: the debate continues. *Curr Opin Rheumatol* 2017;29:317–22.
34 Lukas C, Cyteval C, Dougados M, *et al*. MRI for diagnosis of axial spondyloarthritis: major advance with critical limitations 'Not everything that glisters is gold (standard)'. *RMD Open* 2018;4:e000586.
35 Sepriano A, Landewé R, van der Heijde D, *et al*. Predictive validity of the ASAS classification criteria for axial and peripheral spondyloarthritis after follow-up in the ASAS cohort: a final analysis. *Ann Rheum Dis* 2016;75:1034–42.
36 Baraliakos X, Braun J. Non-radiographic axial spondyloarthritis and ankylosing spondylitis: what are the similarities and differences?: Table 1. *RMD Open* 2015;1:e000053.
37 Sieper J, Hu X, Black CM, *et al*. Systematic review of clinical, humanistic, and economic outcome comparisons between radiographic and non-radiographic axial spondyloarthritis. *Semin Arthritis Rheum* 2017;46:746–53.
38 Costantino F, Aegerter P, Dougados M, *et al*. Two phenotypes are identified by cluster analysis in early inflammatory back pain suggestive of spondyloarthritis: results from the DESIR cohort. *Arthritis & Rheumatology* 2016;68:1660–8.
39 de Winter JJ, Paramarta JE, de Jong HM, *et al*. Peripheral disease contributes significantly to the level of disease activity in axial spondyloarthritis. *RMD Open* 2019;5:e000802.
40 López-Medina C, Dougados M, Ruyssen-Witrand A, *et al*. Evaluation of concomitant peripheral arthritis in patients with recent onset axial spondyloarthritis: 5-year results from the DESIR cohort. *Arthritis Res Ther* 2019;21:139.
41 López-Medina C, Moltó A, Dougados M. Peripheral manifestations in spondyloarthritis and their effect: an ancillary analysis of the ASAS-COMOSPA study. *J Rheumatol* 2019:jrheum.181331.
42 van Lunteren M, van der Heijde D, Sepriano A, *et al*. Is a positive family history of spondyloarthritis relevant for diagnosing axial spondyloarthritis once HLA-B27 status is known? *Rheumatology* 2019.
43 de Hooge M, van Gaalen FA, Renson T, *et al*. Low specificity but high sensitivity of inflammatory back pain criteria in rheumatology settings in Europe: confirmation of findings from a German cohort study. *Ann Rheum Dis* 2019.
44 Poddubnyy D, Callhoff J, Spiller I, et al. diagnostic accuracy of inflammatory back pain for axial spondyloarthritis in rheumatological care. *RMD Open* 2018;4:e000825.

CLINICAL SCIENCE

Immune checkpoint inhibitor-induced inflammatory arthritis persists after immunotherapy cessation

Tawnie J Braaten,[1] Julie R Brahmer,[2] Patrick M Forde,[2] Dung Le,[2] Evan J Lipson,[2] Jarushka Naidoo,[2] Megan Schollenberger,[2] Lei Zheng,[2] Clifton O Bingham 3rd,[1] Ami A Shah,[1] Laura C Cappelli [1]

Handling editor Josef S Smolen

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-216109).

[1]Rheumatology, Johns Hopkins School of Medicine, Baltimore, Maryland, USA
[2]Oncology, Johns Hopkins School of Medicine, Baltimore, Maryland, USA

Correspondence to
Dr Laura C Cappelli, Rheumatology, Johns Hopkins School of Medicine, Baltimore, MD 21224, USA; lcappel1@jhmi.edu

AAS and LCC contributed equally.

Received 1 August 2019
Revised 10 September 2019
Accepted 12 September 2019
Published Online First 20 September 2019

ABSTRACT

Objective We sought to investigate the long-term outcomes of patients who develop immune checkpoint inhibitor (ICI)-induced inflammatory arthritis (IA), to define factors associated with IA persistence after ICI cessation, the need for immunosuppressants and the impact of these medications on underlying malignancies.

Methods We conducted a prospective observational study of patients referred for IA associated with ICIs. Patients were recruited from June 2015 to December 2018. Information was obtained at the baseline visit, and follow-up visits occurred at varying intervals for up to 24 months from ICI cessation. Kaplan-Meier curves were developed to characterise IA persistence. Cox proportional hazards models were used to assess the influence of various factors on IA persistence. Logistic regression was used to evaluate the impact of IA treatment on tumour response.

Results Sixty patients were monitored with a median follow-up after ICI cessation of 9 months. A majority (53.3%) had active IA at their most recent follow-up. IA was less likely to improve in those with longer duration of ICI use, in those receiving combination ICI therapy, and in patients with multiple other immune-related adverse events. Tumour response did not appear to be impacted by immunosuppression. Although not statistically significant, persistent IA was correlated with a better tumour response (complete or partial response).

Conclusion ICI-induced IA can become a long-term disease necessitating management by rheumatology for immunomodulatory treatment. Importantly, the use of immunomodulatory treatment has not been shown to impact cancer outcomes in this study.

Key messages

- Immune checkpoint inhibitor (ICI)-induced inflammatory arthritis (IA) may persist after ICI cessation.
- Longer ICI exposure, receipt of combination ICI therapy and a history of other immune-related adverse events increase the risk of IA persistence.
- Immunomodulatory treatments were efficacious for symptom control while having no apparent effect on tumour response at follow-up.
- Persistent arthritis may be associated with better tumour response (complete or partial response).

INTRODUCTION

The use of immune checkpoint inhibitor (ICI) therapy is rapidly growing, and new agents continue to be approved by the US Food and Drug Administration and European Medicines Agency.[1 2] An anti-CTLA4 agent, ipilimumab, was first approved for metastatic melanoma in 2011. Since then, approvals have been granted for: antiprogrammed death receptor-1 (PD-1) and anti-PD ligand-1 (PDL-1) medications.[3] These agents are approved to treat an increasingly wide variety of cancers, both in the metastatic and adjuvant settings.[1]

While ICIs have improved overall survival and cancer progression-free duration, they can be associated with various autoimmune and inflammatory syndromes known as immune-related adverse events (irAEs).[4] These events are thought to occur through the non-specific activation of T-cells against 'self' antigens.[5] Some irAEs, such as colitis and pneumonitis, can be life-threatening, while others such as inflammatory arthritis (IA) can have a dramatic impact on quality of life. IA is often under-recognised likely due to its limited impact on mortality, examination findings that may miss the threshold of detection by oncology providers, and a broad range of classification options in the Common Terminology Criteria for Adverse Events grading used in clinical trials. However, the necessity for early recognition of IA is growing due to patient functional loss, reports of rapid development of erosions and persistence of joint symptoms.[6–10] The true incidence of IA resulting from irAEs is difficult to determine, but up to 43% of patients in immunotherapy clinical trials reported arthralgias, and it has been estimated that between 3.0% and 7.5% of patients treated with ICIs develop IA.[11–14] In light of the percentage of patients who develop IA and the growing use of ICI therapy, further evaluation of long-term outcomes is warranted, particularly in the setting of improved survival on ICI treatment and previous observations that IA persists after ICI cessation.[7–10]

Little is known about risk factors for persistent IA, appropriate treatment strategies in these patients, or the potential impact of immunosuppressive therapy on tumour response. The treatment of IA is challenging in that individuals may require immunosuppression but have recently received an immune-activating agent. Previous studies note that a majority of patients required

Linked

► http://dx.doi.org/10.1136/annrheumdis-2019-216510

Check for updates

To cite: Braaten TJ, Brahmer JR, Forde PM, *et al*. *Ann Rheum Dis* 2020;**79**:332–338.

systemic corticosteroids for treatment of IA, and of those 15%–90% require additional immunosuppression with conventional synthetic (CS) disease-modifying antirheumatic drugs (DMARDs) or biological (b) DMARDs including tumour necrosis factor (TNF)-inhibitors and interleukin-6 inhibition.[8 9 15–17] One study reported no impact of non-corticosteroid immunosuppression on existing antitumour responses in patients with ICI-related IA.[8]

This study investigated the long-term outcomes of patients who developed IA associated with ICI therapy. We focused on those who developed persistent symptoms after ICI cessation and further evaluated whether factors such as combination ICI therapy, ICI exposure duration or the development of other irAEs were predictive of ongoing IA. Lastly, we assessed if immunosuppressive treatment for IA had an impact on the underlying tumour response to ICIs.

MATERIALS AND METHODS

Patients

This is a prospective observational study of patients referred to the Johns Hopkins Arthritis Center for IA due to ICIs. Patients were recruited from June 2015 to December 2018. Patients were included if they were age 18 or older, had received treatment with an ICI (anti-CTLA4, anti-PD1, anti-PDL-1 or combination ICI therapy) and had a rheumatologist-confirmed diagnosis of IA. We excluded patients if they had a pre-existing rheumatological autoimmune disease, were actively receiving ICI therapy or restarted ICI therapy during follow-up, or if they were enrolled in a trial involving investigational agents for which results were not published.

Patient involvement

Patients with ICI-induced IA were not formally involved in the design for this study. However, discussions with patients by the authors emphasised the importance to patients of understanding whether IA would persist and how likely this would be, the key questions of this study.

Clinical measures

Baseline data including demographics, cancer type, specific ICI treatment, tumour response, personal and family history of autoimmunity, other irAEs, labs (rheumatoid factor (RF), anticyclic citrullinated peptide (CCP), antinuclear antibody (ANA), erythrocyte sedimentation rate (ESR), C reactive protein (CRP) and clinical examination findings (28-tender joint count, 28-swollen joint count, dactylitis and/or enthesitis) were obtained at the initial visit. Follow-up occurred at varying intervals for up to 24 months after ICI cessation to assess the clinical status of IA and malignancy. IrAE information was obtained through medical records review. Tumour response was determined by Response Evaluate Criteria in Solid Tumours 1.1 (RECIST) when available.[18] Otherwise, tumour response was as documented by the patient's treating oncologist in the medical record (ie, in the assessment portion of the progress note or imaging reports). Tumour response was recorded at follow-up visits. Active arthritis was defined as the presence of joint disease (synovitis, tender joints, enthesitis and dactylitis) based on rheumatologist examination or the inability to taper immunosuppressive therapy without return of IA symptoms. If patients had no evidence of active IA at around 6 months after ICI cessation, DMARD medications were tapered. Patients were treated with non-steroidal anti-inflammatory drugs (NSAIDs), intra-articular steroids, systemic steroids, csDMARDS (methotrexate (MTX), leflunomide, sulfasalazine, hydroxychloroquine (HCQ) and bDMARDs (infliximab, adalimumab and etanercept).

Statistics

Descriptive statistics were calculated for baseline characteristics. The distributions of variables were analysed; median values are presented as they were more representative of the distribution of data for the cohort. Kaplan-Meier curves were used to evaluate the persistence of arthritis over time and the influence of various features on persistent arthritis. The time origin for survival analysis was the date of ICI cessation with follow-up occurring over 24 months from ICI cessation. The event of interest was resolution of IA. Log rank testing was performed to evaluate for significant differences when comparing survival curves. Univariable and multivariable Cox proportional hazards regression was performed to identify factors that associated with IA persistence with the outcome being time to IA resolution. Due to the large number of tumour types observed, only the most common, melanoma and non-small cell lung cancer (NSCLC), were assessed as covariates in the Cox proportional hazards model. Variables with less than three events were not examined as covariates in the multivariable model. Factors that achieved a $p<0.10$ were included in the multivariable model. Logistic regression analysis was also performed to evaluate the impact of immunosuppression on tumour progression. HRs and ORs were estimated with their 95% CIs. Statistical significance was set at a $p<0.05$. All statistical analyses were performed using Stata/IC software V.15.0.

RESULTS

Baseline patient characteristics

A total of 60 patients, 32 female and 28 male, were monitored with a median follow-up of 9 months and average follow-up of 12 months (range 1–24 months) after ICI cessation (table 1). The median age was 58.5 years. Of the 60 patients, two had a personal history of single organ autoimmune disease (psoriasis, hypothyroidism) and seven had family history of autoimmune disease including rheumatoid arthritis, Crohn's disease, juvenile idiopathic arthritis, polymyalgia rheumatica and ankylosing spondylitis. The patients had a wide range of cancer diagnoses with melanoma being the most common (35%), followed by NSCLC (23%). Gastrointestinal cancers made up 12% of the study population including pancreatic adenocarcinoma, colon adenocarcinoma, hepatocellular carcinoma and duodenal adenocarcinoma. Four patients had genitourinary cancers (renal cell carcinoma, prostate adenocarcinoma and urothelial carcinoma) and three patients had cervical squamous cell carcinoma, clear cell endometrial carcinoma or endometrial carcinosarcoma. Other cancer diagnoses included Hodgkin's lymphoma, Kaposi sarcoma, mesothelioma, mycosis fungoides, neuroendocrine carcinoma, ependymoma, ductal carcinoma of the breast, basal cell carcinoma and cutaneous squamous cell carcinoma. Combination therapy (anti-CTLA-4+anti-PD-1) was used in 30% of patients whereas monotherapy with anti-CTLA-4, anti-PD-1 or anti-PDL-1 was used in 70% of patients. Some patients received multiple agents over time. ICI treatment was stopped for disease progression, treatment completion and/or severe irAEs. Fourteen (23%) had a complete response based on RECIST, imaging reports or treating oncologist documentation.

Baseline features of IA and other irAEs

The median baseline 28 swollen joint count was 6, and the median baseline 28 tender joint count was 2 (table 2). Median

Table 1 Baseline demographics

	Values*, n=60
Demographics	
Female sex, (%)	32 (53.3)
Age (years)	58.5 (52, 68)
BMI (kg/m^2)	26.6 (22.6, 32.6)
Race, (%)	White: 54 (90) Black: 2 (3.3) Asian: 2 (3.3) White/Asian: 1 (1.7) White/Pacific Islander: 1 (1.7)
Ethnicity, (%)	Non-Hispanic: 60 (100)
Personal history of autoimmune disease, (%)	Psoriasis: 1 (1.7) Hypothyroidism: 1 (1.7)
Family history of autoimmune disease, (%)	Crohn's disease: 2 (3.3) Rheumatoid arthritis: 2 (3.3) Ankylosing spondylitis and Crohn's disease: 1 (1.7) Juvenile idiopathic arthritis: 1 (1.7) Polymyalgia rheumatica: 1 (1.7)
Total ICI duration (months)	7 (2, 13)
Follow-up from ICI cessation (months)	9 (5, 20.5)
Tumour type, (%)†	Melanoma: 21 (35) NSCLC: 14 (23.3) Other: 11 (18.3) Gastrointestinal: 7 (11.7) Genitourinary: 4 (6.7) Gynaecologic: 3 (5)
Immunotherapy, (%)	Combination: 18/60 (30) Monotherapy: 42/60 (70)
Baseline tumour response, (%)	CR: 14 (23.3) PR: 10 (16.7) NED: 4 (6.7) SD: 17 (28.3) PD: 15 (25.0)

*Values expressed as number (percentage) or median (IQR) as appropriate.
†Gastrointestinal cancers included pancreatic adenocarcinoma, colon adenocarcinoma, hepatocellular carcinoma and duodenal adenocarcinoma. Genitourinary cancers included renal cell carcinoma, prostate adenocarcinoma and urothelial carcinoma. Gynaecological cancers included cervical squamous cell carcinoma, clear cell endometrial carcinoma, and endometrial carcinosarcoma. Other cancers included Hodgkin's lymphoma, Kaposi sarcoma, mesothelioma, mycosis fungoides, neuroendocrine carcinoma, ependymoma, ductal carcinoma of the breast, basal cell carcinoma and cutaneous squamous cell carcinoma.
BMI, body mass index; CR, complete response; ICI, immune checkpoint inhibitor; NED, no evidence of disease; NSCLC, non-small cell lung cancer; PD, progressive disease; PR, partial response; SD, stable disease.

Table 2 Baseline features of inflammatory arthritis (IA) and other irAEs

	Values*, n=60	Range/titers
Laboratories		
RF Positive, n=56, (%)	1 (1.8)	0–152
CCP Positive, n=55, (%)	3 (5.5)	37, 43, 2777
ANA positive, n=56, (%)	8 (14.3)	No titer—1 1:40–2 1:80–2 1:160–2 1:640–1
ESR, n=53	29 (9, 53)	1–120
CRP (mg/dL), n=55	1.3 (0.2, 5.4)	0.1–15.7
Examination, (%)		
Dactylitis	2 (3.3)	
Enthesitis	3 (5.0)	
SJC[33]	6 (3, 11)	0–24
TJC[33]	2 (1, 4.5)	0–28
Patient global, n=44	40 (20, 72.5)	0–100
MD global, n=58	27.5 (15, 40)	0–80
CDAI, n=45	17.5 (12, 23)	3–56
Baseline arthritis medications, (%)	Infliximab (for colitis), MTX, and steroid: 1 (1.7) HCQ: 1 (1.7) Steroid alone: 18/60 (30) Baseline NSAIDs: 18/60 (30)	
Baseline prednisone equivalent dose (mg)	10 (5, 50)	2.5–180
Other irAEs, (%)	30 (50.0)	
Specific irAEs, n=30, (%)	Rash: 10 (33.3) Colitis: 10 (33.3) Thyroid: 8 (26.7) Sicca: 7 (23.3) Pneumonitis: 6 (20)‡ Hepatitis: 4 (13.3) Hypophysitis: 4 (13.3) Vitiligo: 3 (10) Pancreatitis: 1 (3.3) Sinusitis: 1 (3.3) Osteitis: 1 (3.3) Arrhythmia with reduced EF: 1 (3.3)	

*Values expressed as number (percentage) or median (IQR) as appropriate.
†An additional two patients developed pneumonitis after their IA visit.
ANA, antinuclear antibody; CCP, cyclic citrullinated peptide antibody; CDAI, clinical disease activity index; CRP, C reactive protein; EF, ejection fraction; ESR, erythrocyte sedimentation rate; HCQ, hydroxychloroquine; irAEs, immune-related adverse events; MTX, methotrexate; NSAID, non-steroidal anti-inflammatory drug; RF, rheumatoid factor; SJC, swollen joint count; TJC, tender joint count.

Clinical Disease Activity Index was 17.5, indicating moderate disease activity. Overall, there were low rates of seropositivity (RF 1.8%, CCP 5.5%, ANA 14.3%) on laboratory analysis. Median baseline ESR and CRP were 29 mm/hour and 1.3 mg/dL respectively. Two patients received csDMARD (MTX, HCQ) for IA prior to their initial evaluation in our clinic. Eighteen (30%) were receiving steroids with a median prednisone dose equivalent of 10 mg daily at the time of presentation to our clinic. NSAIDs alone had been used in 30% of patients prior to initial rheumatological evaluation. Thirty (50%) patients had experienced other non-rheumatic irAEs; rash and colitis were the most common, each in 33% of patients, but a range of other IRAEs were seen (table 2).

Follow-up IA activity

A majority (53.3%) of patients had active arthritis at their last follow-up visit which varied from 1 to 24 months after ICI cessation. Three-month follow-up data after ICI cessation was available in 51 patients, with 6-month data available on 41 patients (online supplementary figure 1). At 3 months, 70.6% had active IA; 48.8% had active IA at 6 months. Among the 20 patients with persistent arthritis at 6 months, 14 continued to have active disease in further follow-up. Kaplan-Meier curves for persistent arthritis showed that those treated with combination immunotherapy and those with ≥2 irAEs were more likely to persist than those treated with monotherapy or those with <2 additional irAEs, respectively (figure 1A,B). The curves evaluating persistent arthritis by category of tumour response did not have a statistically significant difference, but showed a trend towards more persistent arthritis in those with complete or partial tumour responses (figure 1C). In univariate analysis (table 3), arthritis was less likely to improve in those with longer duration of ICI exposure (HR 0.93, 95% CI 0.87 to 0.99;

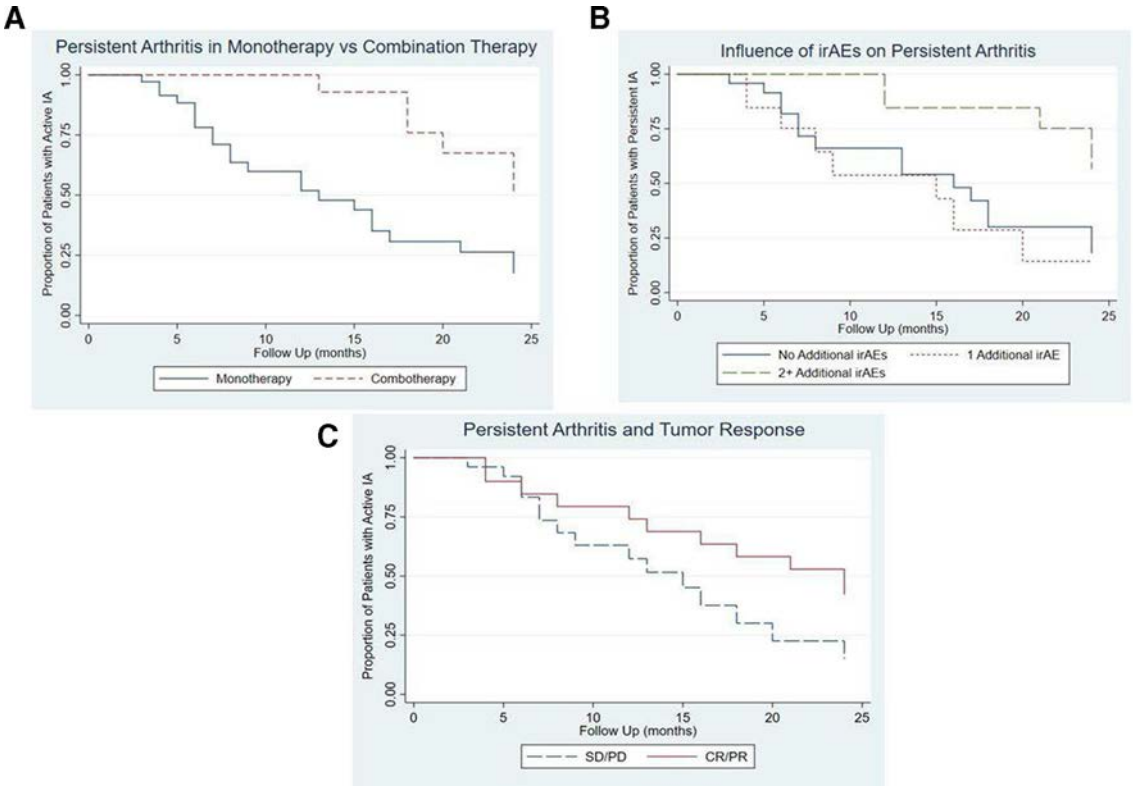


Figure 1 Kaplan-Meier curves for persistent arthritis based on immunotherapy regimen, irAEs, and tumour response. (A) demonstrates increased arthritis persistence in those treated with combination ICI therapy. (B) shows more persistent arthritis in those with >2 additional irAEs. (C) shows that patients with a tumour response of complete response or partial response (CR/PR) had more persistent arthritis compared with those with stable disease or progressive disease (SD/PD) at follow-up. IA, inflammatory arthritis; ICI, immune checkpoint inhibitor; irAEs, immune-related adverse events.

p=0.02), in those receiving combination ICI therapy (HR 0.29, 95% CI 0.12 to 0.72; p=0.008) and in patients with history of other irAEs (HR 0.61, 95% CI 0.39 to 0.95, p=0.03). Duration of ICI treatment (HR 0.82, 95% CI 0.73 to 0.92; p=0.001) and combination ICI therapy (HR 0.06, 95% CI 0.01 to 0.50; p=0.009) remained significant in multivariable analysis (table 4). Although not statistically significant, persistent arthritis was associated with a positive antitumour response rather than stable or progressive disease (HR 0.50, 95% CI 0.22 to 1.11, p=0.09; table 3).

IA treatment and impact on tumour response

Overall, immunomodulatory treatment was required in 75% of patients to treat the arthritis. Forty-eight patients (80%) were treated with systemic and/or intraarticular steroids. csDMARDs were used in 19 patients, and bDMARDS were required in 11 (online supplementary table 1). Of the 24 patients treated with csDMARDs, bDMARDs or combination of both, 4 (16.7%) had progression of their cancer during follow-up evaluation, which was not significantly different than tumour progression in 8 (22.2%) of 36 patients who did not receive DMARDs (OR 0.65, CI 0.17 to 2.47). Of patients who progressed while on DMARDs, one patient was on HCQ, one was on MTX and two were on adalimumab (one for 2 months, the other for 6 months). Adalimumab was stopped in both individuals once cancer progression was noted. In this limited sample size, there was not a statistically significantly increased odd of tumour progression with use of immunosuppressive medications from any category (table 5).

DISCUSSION

ICI-induced IA is an irAE with the potential to become a chronic disease as demonstrated by the persistence of disease at 3 and 6 months after ICI discontinuation. Factors associated with IA persistence included use of combination ICI therapy, longer duration of ICI exposure and prior development of other irAEs. This information provides insight into which individuals are at highest risk for developing persistent IA, thus warranting close monitoring for joint-related symptoms, early referral to and close follow-up by rheumatology and potentially more aggressive immunosuppressive therapy.

That IA can remain active for months to years after ICI cessation is a significant confirmation of previous reports from smaller series,[7–10] especially given the paucity of published data on the persistence of other irAEs. Though there are case reports of colitis relapsing after cessation of ICIs,[19] most gastrointestinal irAEs resolve within 3 months of first symptoms and have minimal risk of recurrence.[20] In contrast, another irAE that may recur or become chronic is pneumonitis; in one cohort, 3 of 19 patients with pneumonitis developed recurrent episodes despite remaining off ICIs.[21] In a recent poster presentation of six patients with persistent irAEs requiring immunosuppression 6 months after ICI cessation, two had arthritis, and one each had hepatitis, colitis, dermatitis and neuropathy.[22] Delayed initial presentation of irAEs, with symptoms starting even after ICI cessation, is a related concept that is being increasingly recognised.[23] Further prospective cohort studies are needed to determine which irAEs besides IA are likely to persist or present after ICI cessation. The types of irAEs that persist may indicate

Table 3 Cox proportional hazards model: univariable analysis to evaluate factors associated with arthritis persistence

	HR	95% CI	P value*
Female sex	0.86	0.40 to 1.85	0.70
Age (years)	1.02	0.99 to 1.05	0.12
BMI (kg/m^2)	0.95	0.88 to 1.02	0.17
Family history of autoimmunity	0.55	0.13 to 2.31	0.41
TJC[33]	0.77	0.65 to 0.92	**0.004**
SJC[33]	1.0	0.92 to 1.09	1.0
Enthesitis	0.65	0.24 to 1.76	0.40
Physician global assessment	0.97	0.95 to 1.00	0.06
Patient global assessment	0.99	0.97 to 1.01	0.30
CDAI	0.95	0.90 to 1.00	0.07
ESR	1.01	0.99 to 1.02	0.29
CRP (mg/dL)	1.02	0.92 to 1.13	0.73
Duration ICI therapy (months)	0.93	0.87 to 0.99	**0.02**
Combo versus Mono therapy	0.29	0.12 to 0.72	**0.008**
ANA positivity	1.77	0.52 to 5.98	0.36
CCP positivity	1.76	0.23 to 13.37	0.58
irAEs (0, 1, 2 or more)	0.61	0.39 to 0.95	**0.03**
Melanoma	0.49	0.22 to 1.12	0.09
NSCLC	0.89	0.36 to 2.20	0.80
Tumour response (CR/PR vs SD/PD)	0.50	0.22 to 1.11	0.09

p-values in bold are statistically significant (<0.05).
*The outcome of interest was resolution of IA. Therefore, HRs <1 denote factors that associate with persistence of IA.
ANA, antinuclear antibody; BMI, body mass index; CCP, cyclic citrullinated peptide antibody; CDAI, Clinical Disease Activity Index; CRP, C reactive protein; CR/PR, complete response and partial response; ESR, erythrocyte sedimentation rate; IA, inflammatory arthritis; ICI, immune checkpoint inhibitor; irAEs, immune-related adverse events; NSCLC, non-small cell lung cancer; SD/PD, stable disease and progressive disease; SJC, swollen joint count; TJC, tender joint count.

how the immune system interacts with particular target tissue microenvironments causing a feed forward loop of autoimmunity that becomes independent of ICIs.

Similar to endocrine-related irAEs, ICI-induced IA may require chronic treatment.[7 8 10 12 16 24] The risk of DMARD treatment is different than hormone replacement due to the concern that immunosuppression may negatively affect antitumour immune responses. In our small sample size, there was not a change in

Table 4 Cox proportional hazards model: multivariable analysis to evaluate factors associated with arthritis persistence

	HR	95% CI	P value*
TJC[33]*	0.88	0.62 to 1.26	0.50
MD global assessment	0.96	0.87 to 1.05	0.33
CDAI	1.03	0.89 to 1.19	0.69
Duration ICI therapy (months)	0.82	0.73 to 0.92	**0.001**
Combo versus monotherapy	0.06	0.01 to 0.50	**0.009**
irAEs (0, 1, 2 or more)	1.05	0.44 to 2.52	0.90
Melanoma	0.41	0.10 to 1.78	0.24
Tumour response (CR/PR vs SD/PD)	1.51	0.34 to 6.63	0.59

Bold values indicate p-values less than 0.05.
*The outcome of interest was resolution of IA. Therefore, HRs <1 denote factors that associate with persistence of IA.
CDAI, Clinical Disease Activity Index; CR/PR, complete response and partial response; IA, inflammatory arthritis; ICI, immune checkpoint inhibitor; irAEs, immune-related adverse events; SD/PD, stable disease and progressive disease; TJC, tender joint count.

Table 5 Univariate analysis: IA treatment and follow-up tumour response (TR)

Medication exposure	# of patients	Worse TR*	OR	95% CI	P value
Systemic steroid alone	20	6	2.54	0.69 to 9.32	0.16
csDMARD	19	2	0.34	0.07 to 1.74	0.20
bDMARD	11	2	0.82	0.15 to 4.43	0.82
Any DMARD	24	4	0.65	0.17 to 2.47	0.53
NSAIDS alone	7	1	0.61	0.07 to 5.58	0.66

*Worse TR: worsened tumour response on follow-up assessment as compared with baseline tumour response (eg, a patient with stable disease at baseline had progression of disease at follow-up).
bDMARD, biological disease-modifying antirheumatic drugs; csDMARD, conventional synthetic disease modifying antirheumatic drugs; IA, inflammatory arthritis; NSAIDs, non-steroidal anti-inflammatory drugs.

tumour response in those treated with csDMARDs or TNF inhibition. This is consistent with previous studies that have shown that treatment of irAEs does not affect tumour response.[8 25]

This study demonstrated that those treated with combination ICI therapy were more likely to have persistent arthritis. The impact of combination ICI therapy versus monotherapy on irAEs has previously been evaluated. Postow *et al* reported that severe irAEs were noted in 54% of patients treated with combination therapy compared with 24% of those treated with anti-CTLA-4 monotherapy for melanoma.[26] Larkin *et al* demonstrated similar findings in melanoma patients with grade 3 or 4 irAEs occurring in 55% of the nivolumab plus ipilimumab group versus 27.3% in the ipilimumab alone group and 16.3% in the nivolumab alone group.[27] While studies have not evaluated whether IA is more likely to develop in those treated with combination ICI therapy versus monotherapy, previous studies have demonstrated that the type of immunotherapy regimen may influence the type of arthritis, with combination therapy associated with more large joint involvement at onset versus more small joint involvement developing in patients treated with anti-PD-1/PD-L1 monotherapy.[8] Further evaluation is warranted to examine why the combination of CTLA-4 and PD-1 blockade is more likely to result in large joint involvement and persistent IA disease.

The impact of duration of immunotherapy on specific irAEs is less well defined, as most irAEs occur during the induction period of ICI treatment.[24] However, various studies have shown that arthritis is a later occurring irAE and can develop months after ICI cessation.[7–10 13–16] We demonstrated that individuals treated with ICI therapy for longer duration were more likely to have persistent arthritis at follow-up. The kinetics of ICI-induced IA development differs from many other irAEs and may suggest unique immune pathogenesis for ICI-induced IA, potentially due to longer exposure to ICI therapy.

Appropriate duration of IA treatment remains unknown. One study noted that IA treatment was continued for an average of 9 months after ICI cessation.[10] In the current report, some patients required immunosuppression >24 months after ICI cessation. Since ICI-induced IA is a novel condition, further evaluation is needed to determine the optimal initial treatment, appropriate dose and taper schedule, and duration of therapy.

Interestingly, our data suggest that the development of persistent arthritis may associate with better antitumour responses compared with patients who have transient IA. This could reflect ongoing activation of the immune system that portends better antitumour immunity. Other studies have demonstrated that development of irAEs associates with better

progression-free and overall survival, but persistence of an irAE has not been examined.[12 28–31] In a study of melanoma patients, patients with irAEs of any grade had improved survival compared with those without irAEs.[29] Studies in non-melanoma patients have also demonstrated similar findings: patients who develop irAEs experience better tumour response.[31 32]

The limitations of this study include a possible selection bias for more symptomatic individuals, as it included only patients referred to rheumatology for joint complaints. Less severe cases may have been self-limited or managed by oncology. Also, patients who had persistent IA may have been more likely to pursue longer rheumatology follow-up. A delay in diagnosis may have occurred in the setting of corticosteroids and/or DMARD therapy to manage other irAEs antecedent to the arthritis. Patients were omitted from analysis if they were on a blinded clinical trial or receiving an investigational immunotherapy agent. Additionally, follow-up was limited in some cases due to death in this patient population. Finally, conclusions about the relationship between tumour response and persistent arthritis are limited as not all patients had RECIST scoring. This heterogeneity in monitoring may have under or overestimated the proportion of patients with a positive tumour response.

This study is one of the largest longitudinal reports to date of patients with ICI-induced IA and the first to evaluate persistence of ICI-induced IA and identify influential factors on outcome. Future studies include evaluating genetic risk factors for development of persistent IA. A previous report found that shared epitope alleles were more common in patients with ICI-induced IA; whether this or other genetic factors play a role in IA chronicity is unknown.[33] Best clinical practices for treating patients with ICI-induced IA based on risk of developing chronic symptoms must also be determined. Those at high risk of developing persistent IA may warrant more specific monitoring approaches by oncologists or pre-emptive treatments such as HCQ. Another area for inquiry is the differences in timing of development and persistence between IA and other irAEs. Understanding specifics in the biology underlying different irAEs can lead to specifically targeted therapy. Overall, continued clinical and translational investigation on larger longitudinal cohorts will allow for increased understanding of pathophysiology and determination of the best clinical care for patients with ICI-induced IA.

Correction notice This article has been corrected since it published Online First. The author names have been amended.

Acknowledgements We acknowledge Michelle Jones for her contributions as a research coordinator for this manuscript.

Contributors LCC, TJB, AAS and COB III designed the study, acquired and analysed data. All authors were involved in interpretation of data. All authors drafted and edited the manuscript.

Funding The database supporting this study was funded in part through a grant from Bristol-Myers Squibb. Additional support was received from an arthritis fellowship award from AbbVie. Additional support was received from the Camille Julia Morgan Arthritis Research and Education fund, the Jerome L. Greene Foundation, and Core B of P30-AR070354 from the National Institutes of Health (NIH).

Disclaimer These sponsors had no role in study design, data analysis or the decision to publish.

Competing interests Individual sources of potential conflict are as follows: JRB: Consultant/Advisory Boards: Bristol-Myers Squibb, Celgene, Lily Merck, Genentech, Amgen, Janssen, Syndax. Grant/Research Funding: Bristol-Myers Squibb, Merck, MedImmune/AstraZeneca. PMF: Consultant/Advisory Board: Abbvie AstraZeneca, Boehringer, Bristol-Myers Squibb, Celgene, Eli Lilly, Novartis. Grant/Research Funding: AstraZeneca, BMS, Corvus, Kyowa, Novartis. DL: Consultant/Advisory Board: Bristol-Myers Squibb, Merck. Grant/Research Funding: Bristol-Myers Squibb, Merck, Aduro Biotech. Honoraria: Merck. EJL: Consultant/Advisory Board: Bristol-Myers Squibb, Novartis, EMD Serono, Array BioPharma, Macrogenics, Merck, Millennium. Grant/Research Funding: Bristol-Myers Squibb, Merck, Sysmex. Patent/Royalty/Intellectual Property: Method of prevent organ transplant rejection using agonist to the PD-1 checkpoint pathway. JN: Consulting/Advisory Board: AstraZeneca, Roche/Genentech, Bristol-Myers Squibb. Grant/Research Funding: Merck, AstraZeneca. Honoraria: Bristol-Myers Squibb, AstraZeneca. LZ: Consultant/Advisory Boards: Biosynergics, Alphamab, Mingrui, Foundation Medicine, NovaRock Biological, Datareve. Grant/Research Funding: Halozyme, iTeos, Bristol-Myers Squibb, Merck, Amgen, NovaRock. Licensing Agreement: Aduro Biotech (given up personal royalty). COB III: Consultant/Advisory Board: Bristol-Myers Squibb, Genentech/Roche, Regeneron/Sanofi. Grant/Research Funding: Bristol-Myers Squibb. AAS: Consultant/Advisory Board: Bristol-Myers Squibb. LCC: Consultant/Advisory Board: Regeneron/Sanofi. Grant/Research Funding: Bristol-Myers Squibb.

Patient consent for publication Not required.

Ethics approval This study was approved by the Johns Hopkins Institutional Review Board IRB00123172.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

ORCID iD

Laura C Cappelli http://orcid.org/0000-0003-2795-7059

REFERENCES

1 Timeline of Progress in Immunotherapy - Cancer Research Institute. Available: https://www.cancerresearch.org/immunotherapy/timeline-of-progress
2 EMA: European Medical Agency. Available: ema.europa.eu/en
3 The Department of Health and Human Services. Fda approves first treatment for advanced form of the second most common skin cancer, 2018. Available: https://www.fda.gov/news-events/press-announcements/fda-approves-first-treatment-advanced-form-second-most-common-skin-cancer-0
4 Weber JS, Kähler KC, Hauschild A. Management of immune-related adverse events and kinetics of response with ipilimumab. *JCO* 2012;30:2691–7.
5 Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer* 2012;12:252–64.
6 Albayda J, Dein E, Shah AA, *et al*. Sonographic findings in inflammatory arthritis secondary to immune checkpoint inhibition: a case series. *ACR Open Rheumatology* 2019;1:303–7.
7 Cappelli LC, Gutierrez AK, Baer AN, *et al*. Inflammatory arthritis and sicca syndrome induced by nivolumab and ipilimumab. *Ann Rheum Dis* 2017;76:43–50.
8 Cappelli LC, Brahmer JR, Forde PM, *et al*. Clinical presentation of immune checkpoint inhibitor-induced inflammatory arthritis differs by immunotherapy regimen. *Semin Arthritis Rheu* 2018.
9 Calabrese C, Kirchner E, Kontzias K, *et al*. Rheumatic immune-related adverse events of checkpoint therapy for cancer: case series of a new nosological entity. *RMD Open* 2017;3:e000412.
10 Smith MH, Bass AR. Arthritis after cancer immunotherapy: symptom duration and treatment response. *Arthritis Care Res* 2019;71:362–6.
11 Cappelli LC, Gutierrez AK, Bingham CO, *et al*. Rheumatic and musculoskeletal immune-related adverse events due to immune checkpoint inhibitors: a systematic review of the literature. *Arthritis Care Res* 2017;69:1751–63.
12 Kostine M, Rouxel L, Barnetche T, *et al*. Rheumatic disorders associated with immune checkpoint inhibitors in patients with cancer—clinical aspects and relationship with tumour response: a single-centre prospective cohort study. *Ann Rheum Dis* 2018;77:393–8.
13 Buder-Bakhaya K, Benesova K, Schulz C, *et al*. Characterization of arthralgia induced by PD-1 antibody treatment in patients with metastasized cutaneous malignancies. *Cancer Immunol Immunother* 2018;67:175–82.
14 Lidar M, Giat E, Garelick D, *et al*. Rheumatic manifestations among cancer patients treated with immune checkpoint inhibitors. *Autoimmun Rev* 2018;17:284–9.
15 Richter MD, Crowson C, Kottschade LA, *et al*. Rheumatic syndromes associated with immune checkpoint inhibitors: a single-center cohort of Sixty-One patients. *Arthritis Rheumatol* 2019;71:468–75.
16 Mooradian MJ, Nasrallah M, Gainor JF, *et al*. Musculoskeletal rheumatic complications of immune checkpoint inhibitor therapy: a single center experience. *Semin Arthritis Rheu* 2018.
17 Kim S, Tayar J, Trinh V, *et al*. Successful treatment of arthritis induced by checkpoint inhibitors with tocilizumab: a case series. *Ann Rheum Dis* 2016;76:2061–4.
18 Eisenhauer EA, Therasse P, Bogaerts J, *et al*. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009;45:228–47.
19 Kikuchi H, Sakuraba H, Akemoto Y, *et al*. A case of nivolumab-associated colitis, which relapsed after mucosal healing and was then successfully treated with mesalazine. *Immunological Medicine* 2019;42:39–44.
20 de Malet A, Antoni G, Collins M, *et al*. Evolution and recurrence of gastrointestinal immune-related adverse events induced by immune checkpoint inhibitors. *Eur J Cancer* 2019;106:106–14.
21 Asher N, Marom EM, Ben-Betzalel G, *et al*. Recurrent pneumonitis in patients with melanoma treated with immune checkpoint inhibitors. *Oncologist* 2019;24:640–7.

22 Giles CR, Barron C, McWhirter E. Persistent immune-related adverse events after immune checkpoint inhibitor therapy. poster presentation CAMO 2019 annual scientific meeting, 2019. Available: https://www.researchgate.net/publication/332929394_Persistent_immune-related_adverse_events_after_immune_checkpoint_inhibitor_therapy [Accessed 8/19/2019].
23 Couey MA, Bell RB, Patel AA, *et al*. Delayed immune-related events (dire) after discontinuation of immunotherapy: diagnostic hazard of autoimmunity at a distance. *J Immunother Cancer* 2019;7.
24 Weber JS, Dummer R, de Pril V, *et al*. Patterns of onset and resolution of immune-related adverse events of special interest with ipilimumab. *Cancer* 2013;119:1675–82.
25 Horvat TZ, Adel NG, Dang T-O, *et al*. Immune-Related adverse events, need for systemic immunosuppression, and effects on survival and time to treatment failure in patients with melanoma treated with ipilimumab at Memorial Sloan Kettering cancer center. *JCO* 2015;33:3193–8.
26 Postow MA, Chesney J, Pavlick AC, *et al*. Nivolumab and ipilimumab versus ipilimumab in untreated melanoma. *N Engl J Med Overseas Ed* 2015;372:2006–17.
27 Larkin J, Chiarion-Sileni V, Gonzalez R, *et al*. Combined nivolumab and ipilimumab or monotherapy in untreated melanoma. *N Engl J Med* 2015;373:23–34.
28 Downey SG, Klapper JA, Smith FO, *et al*. Prognostic factors related to clinical response in patients with metastatic melanoma treated by CTL-Associated antigen-4 blockade. *Clinical Cancer Research* 2007;13:6681–8.
29 Freeman-Keller M, Kim Y, Cronin H, *et al*. Nivolumab in resected and unresectable metastatic melanoma: characteristics of immune-related adverse events and association with outcomes. *Clinical Cancer Research* 2016;22:886–94.
30 Teraoka S, Fujimoto D, Morimoto T, *et al*. Early Immune-Related Adverse Events and Association with Outcome in Advanced Non-Small Cell Lung Cancer Patients Treated with Nivolumab: A Prospective Cohort Study. *J Thorac Oncol* 2017;12:1798–805.
31 Haratani K, Hayashi H, Chiba Y, *et al*. Association of immune-related adverse events with nivolumab efficacy in Non–Small-Cell lung cancer. *Jama Oncol* 2017.
32 Judd J, Zibelman M, Handorf E, *et al*. Immune-Related adverse events as a biomarker in Non-Melanoma patients treated with programmed cell death 1 inhibitors. *Oncol* 2017;22:1232–7.
33 Cappelli LC, Dorak MT, Bettinotti MP, *et al*. Association of *HLA-DRB1* shared epitope alleles and immune checkpoint inhibitor-induced inflammatory arthritis. *Rheumatology* 2019;58:476–80.

CLINICAL SCIENCE

Withdrawal of low-dose prednisone in SLE patients with a clinically quiescent disease for more than 1 year: a randomised clinical trial

Alexis Mathian ,[1] Micheline Pha,[1] Julien Haroche,[1] Fleur Cohen-Aubart,[1] Miguel Hié,[1] Marc Pineton de Chambrun,[1] Thi Huong Du Boutin,[1] Makoto Miyara,[2] Guy Gorochov,[2] Hans Yssel,[2] Patrick Cherin,[1] Hervé Devilliers,[3] Zahir Amoura[1]

Handling editor Josef S Smolen

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-216303).

For numbered affiliations see end of article.

Correspondence to
Pr Zahir Amoura;
zahir.amoura@aphp.fr

AM and MP contributed equally.

Received 11 September 2019
Revised 12 November 2019
Accepted 18 November 2019
Published Online First
18 December 2019

ABSTRACT

Objectives To compare the efficacy to prevent flares of maintenance versus withdrawal of 5 mg/day prednisone in systemic lupus erythematosus (SLE) patients with clinically quiescent disease.

Methods A monocentric, 12-month, superiority, open-label, randomised (1:1) controlled trial was conducted with 61 patients continuing 5 mg/day prednisone and 63 stopping it. Eligibility criteria were SLE patients who, during the year preceding the inclusion, had a clinically inactive disease and a stable SLE treatment including 5 mg/day prednisone. The primary endpoint was the proportion of patient experiencing a flare defined with the SELENA-SLEDAI flare index (SFI) at 52 weeks. Secondary endpoints included time to flare, flare severity according to SFI and British Isles Lupus Assessment Group (BILAG) index and increase in the Systemic Lupus International Collaborating Clinics (SLICC) damage index (SDI).

Results Proportion of patients experiencing a flare was significantly lower in the maintenance group as compared with the withdrawal group (4 patients vs 17; RR 0.2 (95% CI 0.1 to 0.7), p=0.003). Maintenance of 5 mg prednisone was superior with respect to time to first flare (HR 0.2; 95% CI 0.1 to 0.6, p=0.002), occurrence of mild/moderate flares using the SFI (3 patients vs 12; RR 0.2 (95% CI 0.1 to 0.8), p=0.012) and occurrence of moderate/severe flares using the BILAG index (1 patient vs 8; RR 0.1 (95% CI 0.1 to 0.9), p=0.013). SDI increase and adverse events were similar in the two treatment groups. Subgroup analyses of the primary endpoint by predefined baseline characteristics did not show evidence of a different clinical response.

Conclusion Maintenance of long term 5 mg prednisone in SLE patients with inactive disease prevents relapse.

Trial registration number NCT02558517; Results

Key messages

What is already known about this subject?

- Little data are available on glucocorticoid (GC) withdrawal for patients with systemic lupus erythematosus (SLE) once remission has been achieved.

What does this study add?

- In SLE patients with quiescent disease and stable treatment regimen for at least 1 year, withdrawal of 5 mg of prednisone was associated with a fourfold increase in the risk of flare.
- No worsening of Systemic Lupus International Collaborating Clinics damage index and the GC toxicity index were observed during the study.

How might this impact on clinical practice or future developments?

- There is an interest of continuing 5 mg prednisone at long course to avoid relapse in SLE.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a chronic disease characterised by a fluctuating disease course. Glucocorticoids (GCs) play a central role in the treatment of active SLE but little data are available on GC withdrawal for patients once remission has been achieved.[1–4] Although there is general agreement on the toxicity of GCs and the need to avoid long-term administration of these drugs, a significant proportion of treating physicians prefers to continue low-dose GCs despite clinical remission, particularly if there is a history of major organ involvement such as lupus nephritis or neuropsychiatric SLE.[5] Observations of SLE cohorts reported that between 57% and 86% of patients undergo long-term low-dose GCs treatment.[6 7]

However, although some authors feel that withdrawal of low-dose GC may lead to severe flares, even after very long intervals of complete quiescence; this concept can neither be proved nor refuted based on the literature. Conversely, some studies suggest that sustained low doses of GCs might be harmful and could be associated with the accrual of irreversible organ damage over time.[8–11] Accordingly, in the recent update of European League Against Rheumatism (EULAR) recommendations for the management of SLE during chronic maintenance treatment, GCs should be, when possible, withdrawn.[2]

Due to the lack of experimental evidence to justify long-term use of low-dose GC administration in SLE, we have conducted the CORTICOLUP study, a monocentric, 12-month, superiority, open-label, randomised controlled clinical trial comparing maintenance versus withdrawal of 5 mg/

To cite: Mathian A, Pha M, Haroche J, *et al*. *Ann Rheum Dis* 2020;**79**:339–346.

day prednisone for the prevention of flares in SLE patients with clinically inactive disease and a stable SLE treatment for at least 1 year.

PATIENTS AND METHODS

The study was conducted from January 2014 to April 2018 in the Department of Internal Medicine 2, French National Reference Center for SLE, Pitié-Salpêtrière Hospital, Paris, France. Written informed consent was obtained from all participants.

Patient eligibility, enrolment, randomisation and treatment

Eligible patients were 18 years or older, with a diagnosis of SLE according to the revised American College of Rheumatology classification criteria[12]; a clinically quiescent SLE for at least 1 year defined as: (1) SELENA-SLEDAI score ≤4,[13 14] (2) D or E British Isles Lupus Assessment Group (BILAG) 2004[15 16] scores in all organ systems except for the haematological system, for which a C score due to leucopenia, lymphopenia or isolated positive Coombs' test was tolerated and (3) Physician's Global Assessment=0[13 14] and a treatment regimen including prednisone 5 mg/day. Prednisone, antimalarials and/or immunosuppressive therapy had to be stable for at least one consecutive year before inclusion. Exclusion criteria were patients who were pregnant, who planned a pregnancy and who were unable to sign the informed consent. Prolonged clinically quiescent SLE was defined as a 5-year consecutive period of no clinical signs of disease activity, irrespective of occurrence of leucopenia, SLE treatment and serological activity (presence of anti-double-stranded DNA (anti-dsDNA) antibodies (Abs) and/or low complement).[17]

Eligible subjects were enrolled in the study and computer-randomly (1:1 ratio) assigned to continue prednisone 5 mg/day for 52 weeks or to interrupt intake the day of the beginning of the study (day 0) (http://randoweb.aphp.fr/). Patients assigned to prednisone withdrawal were given hydrocortisone 20 mg/day during 1 month to prevent adrenal failure. Other SLE treatment including antimalarial and/or immunosuppressant therapy remained unmodified during the study except in case of treatment-related side effects or SLE flare requiring treatment modifications. An isolated change in anti-dsDNA or C3, in the absence of clinical manifestations, was not indicative of an increase in SLE treatment.

Outcomes and follow-up

The primary efficacy endpoint was the proportion of patients experiencing a flare using the SELENA-SLEDAI flare index (SFI)[13 14] (see online supplementary text) between randomisation and week 52.

Secondary endpoints included (see online supplementary text): times to flare; proportion of patients experiencing a severe flare or a mild/moderate flare using the SFI at week 52[13 14]; proportion of patients experiencing a flare using the BILAG index at week 52[18 19]; proportion of patients experiencing a severe flare, a moderate flare or a mild flare using the BILAG index at week 52[18 19]; changes in serological activity (anti-dsDNA Abs and C3 levels) during 52 weeks and proportion of patients experiencing an increase in the Systemic Lupus International Collaborating Clinics (SLICC) damage index (SDI) between randomisation and week 52.[20] Severe and moderate flares in the BILAG index were pooled because this type of flare typically requires the prescription of corticosteroids and/or the initiation or increase of immunosuppressant and/or specific drugs in contrast to mild flares that often require symptomatic therapy. The outcomes were adjudicated by a blinded independent committee of the intervention allocation.

Patients were evaluated at baseline and at 3, 6, 9 and 12 months and were asked to contact their physician if they developed symptoms compatible with an SLE flare after which they were promptly examined. At each visit, outcomes and adverse effects were ascertained according to a history of current symptoms and medications (see online supplementary text). Follow-up data were collected until the end of 52 weeks regardless of outcome, even for subjects who had discontinued intervention.

Statistical analysis

The sample size calculation was based on the assumption that in SLE patients with inactive disease and long-term 5 mg prednisone, the risk of relapse was estimated to be 3% and that a 15 points increase in the percentage of the proportion of flare in the prednisone withdrawal group (i.e 18% flare) was considered clinically significant. Under these assumptions, at least 62 patients would have to be assigned to each group in order to have 80% statistical power permitting to conclude that prednisone maintenance was superior to prednisone withdrawing with a two-sided type I error rate of 5%. Primary and secondary endpoints and safety analyses were analysed in the intention-to-treat principle that included all randomised patients. Qualitative variables are expressed as number (%) and quantitative parameters as the mean ± standard deviation (SD) or median (range or quartiles alternatively), as appropriate. Differences between groups of patients were tested using the Mann-Whitney test for continuous data, and Fisher's exact test or the Khi-2 test for categorical data. Times to flare were represented by means of the Kaplan-Meier method and compared using Log-Rank tests. Hazard ratio were obtained using the Cox proportional hazards model. Interaction between treatment effect and prespecified subgroup effect was tested using a logistic regression model. Since the variables pertaining to "disease duration" and "GCs duration" are closely linked, the former was not included in the interaction analysis. All tests were two-sided and $p<0.05$ defined significance statistical analyses were performed using GraphPad Prism, V.5.0 software (GraphPad Software, San Diego, California, USA) and SAS V.9.4 software.

RESULTS

Baseline demographic and disease characteristics

A total of 124 patients were enrolled. Sixty-one and 63 patients were randomised in the prednisone maintenance and the prednisone withdrawal groups, respectively (figure 1). In the prednisone maintenance group, two patients stopped prednisone intake for personal reasons. In the prednisone withdrawal group, four patients restarted prednisone at 5 mg/day, two for personal reasons and two because of pregnancy. All patients completed the study, including its follow-up. Baseline clinical characteristics at randomisation are summarised in table 1. There were no significant differences between the two treatment groups with respect to any of the baseline clinical characteristics, except a significantly higher number of patients under methotrexate in the maintenance group ($p=0.035$) and under mycophenolate mofetil in the prednisone withdrawal group ($p=0.038$). At study entry, all patients were in 'remission on treatment' according to the DORIS consensus[21 22] and in 'remission on corticosteroids' according to the Zen *et al* definition.[23] A total number of 24 (39%) patients in the maintenance group and 32 (51%) patients in the withdrawal group were in prolonged clinically quiescent SLE.

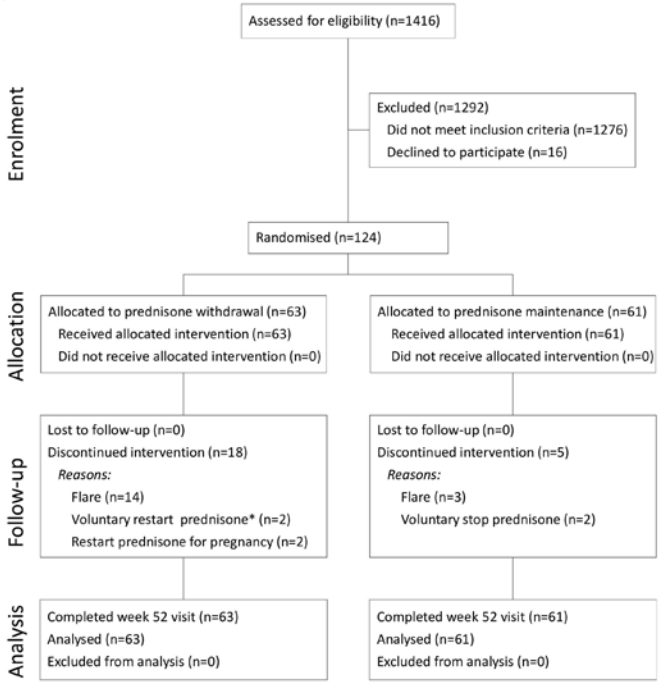


Figure 1 Study flow diagram. In the prednisone withdrawal allocated group, 14 patients experienced a flare that justified restart of prednisone and 4 patients restarted prednisone while not experiencing lupus symptoms (2 for pregnancy and 2 for personal reasons). In the prednisone maintenance allocated group, three patients experienced a flare that justified increase above 5 mg/day of prednisone and two patients stop prednisone 5 mg/day for personal reasons. All patients completed the 52 weeks follow-up and were included in the intention-to-treat analysis. *One patient who restarted prednisone for personal reasons became pregnant afterwards.

Table 1 Baseline demographic and clinical characteristics of the study subjects

Characteristic	Maintenance group (n=61)	Withdrawal group (n=63)
Age, years	41±1.7	44±1.6
Women	55 (90)	56 (89)
Disease duration, months	147±86	163±96
Quiescence duration, months	56±6	68±7
Mean SELENA–SLEDAI score	1.5±0.2	1.4±0.2
History of		
Lupus nephritis	21 (34)	26 (41)
Neuropsychiatric lupus	4 (7)	8 (13)
Arthritis	43 (71)	55 (87)
Cutaneous lupus	35 (57)	39 (62)
Serositis	15 (25)	17 (27)
SDI score	0.5±0.1	0.7±0.2
Low C3	17 (28)	18 (29)
Increased dsDNA binding	29 (48)	28 (46)
Low C3 and increased dsDNA binding	10 (16)	10 (16)
HCQ use	57 (93)	56 (89)
[HCQ], µg/L	1071±66	953±55
[HCQ]>750 µg/L*	38/56† (68)	38/56† (68)
Corticoid duration, months	137±11	145±13
Immunosuppressive drugs	17 (28)	16 (25)
Methotrexate	10 (16)	3 (5)
Azathioprine	3 (5)	1 (2)
Mycophenolate mofetil	4 (7)	12 (19)

Values are expressed as n (%).
Plus–minus values are means±SD.
*Considered to be the therapeutic target.[45]
†Positive assay/number of patients taking HCQ. One HCQ serum concentration was missing in the maintenance group.
C3, Complement fraction 3; dsDNA, double-stranded DNA; [HCQ], blood concentration; HCQ, hydroxychloroquine;SDI, SLICC damage index; SELENA-SLEDAI, Safety of Estrogens in Lupus Erythematosus:National Assessment version of the Systemic Lupus Erythematosus Disease Activity Index; SLICC, Systemic Lupus International Collaborating Clinics.

Withdrawal of low dose of prednisone increases the risk of flares in SLE patients with clinically quiescent disease

Primary endpoint

Using the SFI, the proportion of patients who experienced a flare at 52 weeks differed significantly between the two groups: 4/61 (7%) patients in the maintenance group versus 17/63 (27%) in the withdrawal group (RR 0.2 (95% CI 0.1 to 0.7), p=0.003, see table 2).

Secondary endpoints

SLE flares

The occurrence of flare during 52 weeks is presented as a Kaplan-Meier curve of cumulative probability (figure 2). The estimated hazard ratio (prednisone maintenance/prednisone withdrawal) of flare occurrence was 0.2 (95% CI, 0.1 to 0.6, p=0.002 by log-rank).

When recording flares using the BILAG index, the proportion of patients who experienced a flare at 52 weeks differed significantly between the two groups: 4/61 (7%) patients in the maintenance group versus 17/63 (27%) in the withdrawal group (RR 0.2 (95%CI 0.1 to 0.7), p=0.003, see table 2).

Analysis of flare severity at 52 weeks using the SFI showed a significantly lower proportion of mild/moderate flare (3 patients vs 12; RR 0.2 (95% CI 0.1 to 0.8), p=0.012) and a non-significant lower proportion of severe flare (1 patient vs 5; RR 0.2 (95% CI 0.1 to 1.5), p=0.096) in the maintenance group, as compared with the withdrawal group. Flare severity analysis using the BILAG index assessment showed a significantly lower proportion of moderate/severe flare in the maintenance group as compared with the withdrawal group (1 patient vs 8; RR 0.1 (95% CI 0.1 to 0.9), p=0.013).

The clinical manifestations and treatment of flares are depicted in table 3. There were 4 cases of arthritis, 2 cutaneous manifestations, 1 class V lupus nephritis and 1 mucosal ulcer in the maintenance group and 12 arthritis, 7 cutaneous manifestations, 2 class V lupus nephritis, 1 mucosal ulcer, 1 pericarditis, 1 catatonia, 1 cranial neuropathy and 1 thrombocytopenia in the withdrawal group. In the prednisone maintenance group, three flares were treated with prednisone above 5 mg/day and one with the start of an immunosuppressant drug. In the prednisone withdrawal group, 12 flares were treated with prednisone above 5 mg/day and 4 with a new immunosuppressant or immunomodulatory drug. To be certain that an adrenal insufficiency was not misdiagnosed as a lupus flare, four patients with less evident symptoms of synovitis and arthritis underwent a short Synacthen test that revealed no diagnosis of adrenal insufficiency (online supplementary table S1).

Damage accrual

Four items of the SDI were scored during 52 weeks in a total of three patients in the prednisone withdrawal group: two osteoporosis-related fractures, one retinal toxicity due to antimalarials and one cataract. No damage-related events were

Table 2 Results for primary and secondary endpoints at 52 weeks

	Maintenance group (n=61)	Withdrawal group (n=63)	Relative risk (95% CI)	P value*
Primary endpoint: any flare according to SFI	4 (7)	17 (27)	0.2 (0.1 to 0.7)	0.003
Secondary endpoints:				
Detail of SFI flares				
No flare	57 (93)	46 (73)	1 (Ref.)	Ref.
Mild/moderate	3 (5)	12 (19)	0.2 (0.1 to 0.8)†	0.012
Severe	1 (2)	5 (8)	0.2 (0.1 to 1.5)†	0.096
Any flare according to BILAG index	4 (7)	17 (27)	0.2 (0.1 to 0.7)	0.003
Detail of BILAG index flares				
No flare	57 (93)	46 (73)	1 (Ref.)	Ref.
Mild	3 (5)	9 (14)	0.3 (0.1 to 1.1)†	0.066
Moderate	0 (0)	3 (5)	NA†	0.096
Severe	1 (2)	5 (8)	0.2 (0.1 to 1.5)†	0.096
Moderate/severe	1 (2)	8 (13)	0.1 (0.1 to 0.9)	0.013
Patients experiencing an increase in the SDI	0 (0)	3 (5)	NA	0.244

Values are the number (%) of patients.
*Using Fisher's exact test.
†Compared with patients in the 'no flare' group as the reference.
BILAG, British Isles Lupus Assessment Group; NA, not available; SDI, SLICC damage index; SFI, SELENA-SLEDAI flare index; SLICC, Systemic Lupus International Collaborating Clinics.

recorded in the prednisone maintenance group. The proportion of patients who experienced an increase in the SDI was similar between the two groups (table 2).

Changes in immunological parameters

The anti-dsDNA Ab and C3 serum levels, as well as the proportion of patients with positive anti-dsDNA Abs and those with low C3, did not significantly change during 52 weeks in either group (online supplementary figure S1 and data not shown).

Risk of flare in patients' subgroups

The effect of treatment maintenance was consistent among prespecified subgroups (figure 3).

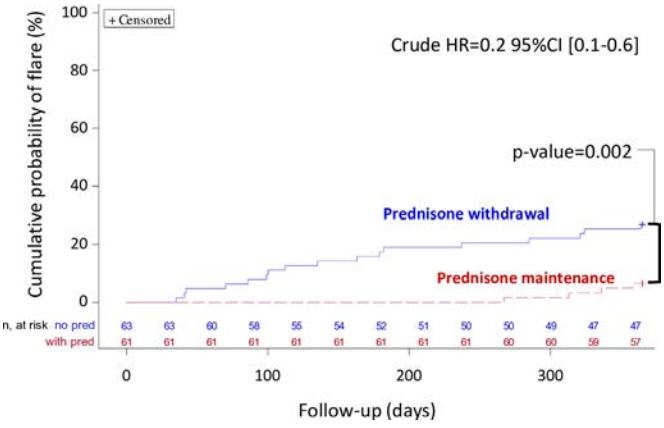


Figure 2 Kaplan-Meier estimates of the cumulative probability of flare for patients in the prednisone maintenance and prednisone withdrawal groups. Clinically quiescent SLE patients were allocated at day 0 to stop (blue line) or to continue (red line) prednisone 5 mg/day and were followed for 52 weeks. Each corner in the curve represents a lupus flare, defined by SELENA-SLEDAI flare index. Patients who had a flare in any organ system were recorded. Kaplan-Meier plots show the percentage of patients who flared in any organ system. All patients had a 52 weeks survey. No patient was censored before the end of the study. Curves were compared using log-rank tests. Crude HR was calculated using a proportional risk COX model. Pred, prednisone; SLE, systemic lupus erythematosus.

Adverse events

During the 52 weeks of the study, adverse events (table 4) were rare in both groups. There were no deaths, no vascular thrombosis, no malignant neoplasm and no adverse events that required the discontinuation of prednisone or hospital admission. Six patients, three in each group, became pregnant during the study. In a post-hoc statistical analysis, the mean±SD variation of the composite glucocorticoid toxicity index (GTI) between baseline and month 12 was similar in the maintenance (3.3±13.0) and the withdrawal groups (3.7±16.5) (p=0.9). The proportion of patients who experienced a worsening in their composite GTI between baseline and month 12 was also similar in the two groups (23% vs 29%, p=0.5).

DISCUSSION

Understanding the benefits and risks of long-term maintenance of low-dose GCs is an important consideration in the care of SLE. The present study prospectively shows that in SLE patients with quiescent disease and stable treatment regimen for at least 1 year, withdrawal of 5 mg of prednisone was associated with a fourfold increase in the risk of flare, as defined by the SFI or the BILAG index, thereby emphasising the interest of continuing a low dose of prednisone at long course to avoid relapse.

GCs tapering and withdrawal are considered one of the main targets of SLE management, but at present the decision of withdraw GCs is left to the judgement of the treating physician. A recent internet-based survey of 130 clinicians from 30 countries showed that preference of clinicians in treatment reduction in patients with SLE in clinical remission was variable with greater caution in treatment reduction when patients have persistent serological abnormalities and previous major organ involvement.[5] Prednisolone was by far the first medication that physician suggested reducing or withdrawing during remission, irrespective of persistent serological abnormalities, remission duration, minor or major organ involvement and whether prednisolone was used with hydroxychloroquine (HCQ) alone or as part of a regimen also involving HCQ and an immunosuppressant.[5] To our knowledge, the present study is the first randomised controlled trial comparing the risk of relapse after withdrawal of low-dose prednisone in clinically

Table 3 Clinical manifestations and treatment of flares according to treatment group

Patient	Time to flare (days)	Clinical manifestations	Flare according to SFI	SELENA-SLEDAI score	PGA	Flare according to BILAG 2004 index	BILAG 2004 score	Treatment
Prednisone maintenance group								
# 23	313	Arthritis, diffuse alopecia, constitutional signs	Mild/moderate	10	1.1	Mild	1B, 2C	Pred 15 mg/day
# 76	358	Arthritis, nephritis (class V), leucopenia	Severe	13	1.5	Severe	1A, 1B, 1C	Pred 30 mg/day+MMF 2 g/day
# 108	336	Mucosal ulcers, mild arthritis, lymphopenia	Mild/moderate	4	0.4	Mild	3C	No change
# 109	267	Arthritis, diffuse alopecia	Mild/moderate	10	1.0	Mild	1B, 1C	Pred 10 mg/day
Prednisone withdrawal group								
# 4	99	Arthritis, mucosal ulcers, subacute cutaneous lupus, malar rash, lymphopenia	Mild/moderate	8	1.5	Moderate	2B, 1C	Pred 10 mg/day+↗ HCQ 600 mg/day+TLD 100 mg/day
# 19	324	Subacute cutaneous lupus, diffuse alopecia, leucopenia	Mild/moderate	7	0.7	Mild	1B, 2C	↗ HCQ 600 mg/day
# 27	285	Constitutional signs, psychosis (catatonia), nephritis (class V)	Severe	19	2.6	Severe	3A	Pred 10 mg/day+MMF 2 g/day+plasmapheresis
# 40	41	Pericarditis, arthritis	Severe	6	2.1	Severe	1A, 1B	IVMP followed by pred 30 mg/day+MTX 15 mg/week+belimumab
# 43	179	Malar rash, diffuse alopecia	Mild/moderate	4	1.0	Mild	1B, 1C	No change
# 56	42	Subacute cutaneous lupus, arthritis	Mild/moderate	6	1.0	Moderate	2B	Pred 5 mg/day
# 58	237	Arthritis	Severe	4	1.0	Severe	1A	IVMP followed by pred 15 mg/day
# 59	163	Arthritis, diffuse alopecia	Mild/moderate	8	0.9	Mild	1B, 1C	Pred 5 mg/day
# 71	35	Cranial neuropathy (acute hearing loss)	Severe	10	1.4	Severe	1A	IVMP followed by pred 60 mg/day
# 72	364	Arthritis	Mild/moderate	4	0.9	Mild	1B	Pred 10 mg/day+↗ MTX 15 mg/week
# 77	86	Arthritis, diffuse alopecia	Mild/moderate	10	0.9	Mild	1B, 1C	Pred 15 mg/day
# 78	182	Arthritis, lymphopenia	Mild/moderate	6	1.0	Mild	1B, 1C	Pred 15 mg/day
# 82	70	Arthritis	Mild/moderate	6	0.9	Mild	1B	Pred 15 mg/day
# 91	112	Arthritis, rash, diffuse alopecia, thrombocytopenia	Mild/moderate	11	1.0	Moderate	2B, 2C	Pred 10 mg/day
#94	135	Arthritis	Mild/moderate	4	0.9	Mild	1B	Pred 10 mg/day
# 95	321	Nephritis (class V)	Severe	6	1.5	Severe	1A	MMF 2 g/day
#107	100	Arthritis, constitutional signs	Mild/moderate	8	0.9	Mild	1B, 1C	Pred 15 mg/day

Cytopenia was defined as leucopenia <3 g/L, lymphopenia <1 g/L and thrombocytopenia <100 g/L.
↗, increase; BILAG, British Isles Lupus Assessment Group;HCQ, hydroxychloroquine; IVMP, intravenous high-dose methylprednisolone; MMF, mycophenolate mofetil; MTX, methotrexate; PGA, Physician's Global Assessment; pred, prednisone; SELENA-SLEDAI, Safety of Estrogens in Lupus Erythematosus:National Assessment version of the Systemic Lupus Erythematosus Disease Activity Index;SFI, SELENA-SLEDAI flare index; TLD, thalidomide.

quiescent SLE. Galbraith *et al* have undertaken a pilot trial comparing continuation and withdrawal of 7.5 mg of prednisone in SLE patients with a history of previous lupus nephritis but this trial was designed to study the feasibility of a larger trial and not to assess the efficacy and safety of this treatment intervention.[24] However, there is some indirect data that underscore the importance of low-dose GCs for maintaining clinical quiescence. Indeed, it is well known that prolonged complete remission defined by the absence of clinical activity with no use of GCs and immunosuppressant is infrequent in SLE ranging from 2% to 32%.[11 17 23 25–27]

The 27% relapse rate observed in the withdrawal group in our study is in line with the ones recently reported in two recent cohorts.[28 29] Tani *et al* described the longitudinal study of a cohort of 91 SLE Italian patients who attempted GC stopping.[28] A total of 77 patients successfully stopped GC. For those patients who were successfully withdrawn from GC, 18 flares (23%) were recorded after a median follow-up period of about 2 years.[28] Like in our study, 72% of flares were mild. The time period since the last flare was the sole determinant predictor of disease flare identified.[28] Goswami *et al* in a recent observational study preformed in India reported that 21% of patients in remission undergo exacerbation of the disease after GC withdrawal with most of the flares occurring in the first year of follow-up.[29] Unlike our study, many flares were deemed major. Furthermore, Goswami *et al* defined duration of disease, duration of GCs before interruption and second immunosuppressive as independent predictors of flare-free survival,[29] whereas in our prospective randomised study none of the above-mentioned factors was associated with the risk of flare after GC withdrawal. However, we cannot completely rule out certain associations since subgroups analyses were performed only on relatively small size samples. In particular, the results might have been different should the majority of patients have been treated with an immunosuppressant.

In other autoimmune diseases, there is evidence that low-dose long duration of GC use is beneficial. A systematic review and meta-analysis of prednisone withdrawal in

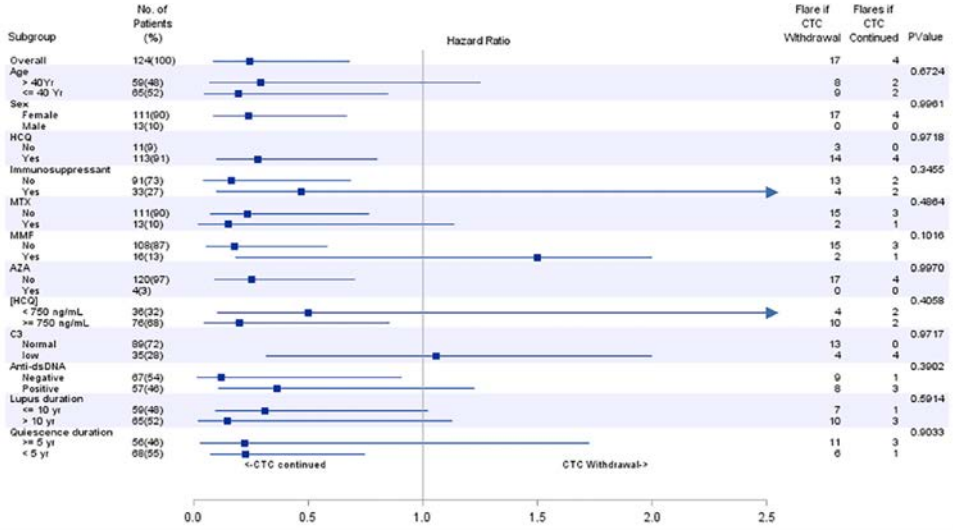


Figure 3 Risk of flare in patients subgroups. The impact of maintenance or withdrawal of prednisone on the occurrence of lupus flare was studied according to predefined subgroups. The HR was derived from a COX model, with treatment as the only factor, according to subgroup. AZA, Azathioprine; C3, complement fraction 3; dsDNA, double-stranded DNA; HCQ, hydroxychloroquine; [HCQ], blood concentration; MMF, mycophenolate mofetil; MTX, methotrexate.

ANCA-associated vasculitis found that late continuation of prednisone was associated with a twofold reduction of the frequency of relapses.[30] Moreover, in rheumatoid arthritis, the results from two randomised, controlled trials have shown that continuation of low-dose GC provided better disease control than GC withdrawal.[31 32] Yet, these results have to be interpreted with caution given the different pathogenesis of these autoimmune diseases.

The interpretation of maintaining low-dose prednisone therapy is subject to prudence, since the CORTICOLUP study has not been able to prove that maintaining low-dose steroids would diminish damage accrual in a statistically significant way. However, the reduction in the number of flares with the maintenance of prednisone at low doses is clearly important because previous reports have shown a strong link between the number of flares and the damage accrual, as well as the quality of life.[33–36] Enthusiasm for long-term prednisone, even if effective, is also tempered by potential side effects such as infections, diabetes mellitus, cataract, osteoporosis, gastrointestinal bleeding and cardiovascular disease, leading to the development of irreversible organ damage.[7 9 10 37–42] Even if some studies suggest that sustained low doses of GCs might be harmful and could be associated with the accrual of irreversible organ damage over time,[8–11] a recent report from the EULAR agreed that the level of harm conveyed by long-term GG administration is dose-dependent with dosages of ≤5 mg/day prednisone producing an acceptably low level of harm, with the exception of patients at high cardiovascular risk who may require preventive measures.[43] In line with these data, we observed a low SDI score on CORTICOLUP inclusion and no worsening of SDI and GTI scores during the study. However, the small number of patients included and the limited monitoring time prevent us from drawing any meaningful conclusion on the tolerance of long-term prednisone maintenance.

Table 4 Adverse events by treatment group

Type of adverse events	Adverse events, n
Prednisone maintenance group	
Pyelonephritis	2
Prednisone withdrawal group	
Osteoporosis-related bone fracture	2
Adrenal insufficiency	1
Pyelonephritis	1
Retinal antimalarials toxicity	1
Cataract	1

Multiple occurrences of the same adverse event in one person were counted only once. SLE flares were not considered adverse events.
SLE, systemic lupus erythematosus .

The results of the study should be interpreted in light of its limitations. First, it was an open-label trial without a placebo group. However, given that outcomes were strictly adjudicated by an independent assessor, they are unlikely to be affected by knowledge of patient allocation. Second, this study was a monocentric investor-led clinical trial: nevertheless, patients were educated to detect signs of flare and contact their physicians and no patients were lost to follow-up. Third, although in this study no data on ethnicity are given, it has to be noted that the French population of SLE patients is widely multi-ethnic, and includes a significant proportion of patients from African and Asian ancestors. Fourth, because withdrawal of 5 mg of prednisone was relatively abrupt, we cannot exclude that slow prednisone tapering would have resulted in less flares. However, there is no experimental data in SLE sustaining this hypothesis. Of note, in a recent randomised, controlled trial with rheumatoid arthritis patients with a low disease activity status, it was found that continued prednisone-equivalent dose 5 mg provided better disease control than GC withdrawal, even if the latter was slowly tapered in monthly 1 mg decrements.[44] This study confirms two previously published reports with a close design.[31 32] Fifth, it can be argued that certain cases of lupus flares in the withdrawal group might have been misdiagnosed as adrenal insufficiency because cortisol levels were not systematically assessed during the study. However, the occurrence of events such as rash, alopecia, arthritis, synovitis, serositis and nephritis cannot be confused with symptoms of adrenal insufficiency. Furthermore, other signs of adrenal insufficiency such as gastrointestinal symptoms, hypoglycaemia, hypotension, hyponatremia, lymphocytosis and

hypereosinophilia were absent as well. Therefore, the clinical and laboratory manifestations observed in patients who have discontinued corticosteroid therapy were clearly those of SLE. Finally, it is possible that our study may suffer from an inclusion bias. The SLE patients were kept on low dose of steroids by their treating physician despite clinical remission. It is possible that these patients have a special lupus history with severe flares, major organ involvements and relapses that prompted the physician to maintain this long-term treatment. Thus one must be circumspect in extrapolating the results to all SLE patients in remission. The results should also be interpreted in the context of approximately three-quarters of patients experiencing GC withdrawal success.

In conclusion, despite some limitations, our study presents to date the strongest evidence that maintenance of 5 mg prednisone is superior to its withdrawal in order to prevent flares in SLE patients in remission. These results must be validated in other independent cohorts and larger studies must be undertaken to determine whether clinical characteristics and new biomarkers, such as elevated serum interferon alpha levels, could help clinicians to identify a subgroup of SLE patients clinically in remission but who are at higher risk of relapse and who would benefit from a continued intake of low doses of prednisone.

Author affiliations
[1]Sorbonne Université, Assistance Publique–Hôpitaux de Paris, Groupement Hospitalier Pitié–Salpêtrière, French National Referral Center for Systemic Lupus Erythematosus, Antiphospholipid Antibody Syndrome and Other Autoimmune Disorders, Service de Médecine Interne 2, Institut E3M, Inserm UMRS, Centre d'Immunologie et des Maladies Infectieuses (CIMI-Paris), Paris, France
[2]Sorbonne Université, Inserm UMR-S, Centre d'Immunologie et des Maladies Infectieuses (CIMI-Paris), Assistance Publique–Hôpitaux de Paris, Groupement Hospitalier Pitié–Salpêtrière, Département d'Immunologie, Paris, France
[3]Centre Hospitalier Universitaire de Dijon, Hôpital François-Mitterrand, service de médecine interne et maladies systémiques (médecine interne 2) et Centre d'Investigation Clinique, Inserm CIC-EC 1432, Dijon, France

Acknowledgements We thank the patients, the nurses and the staff of the Department of Internal Medicine 2 who participated in this study. We thank Professor Anne Bachelot (Sorbonne Université, Assistance Publique–Hôpitaux de Paris, Groupement Hospitalier Pitié–Salpêtrière) for her advices.

Contributors AM, MP, HD and ZA contributed to the conception and design of the study; AM, MP, JH, FC-A, MH, MPdC, THDB, MM, GG, PC and ZA were involved in the acquisition of data; AM, MP, JH, FC-A, MH, MPdC, THDB, MM, GG, HY, PC, HD and ZA contributed to the analysis and interpretation of data. All authors contributed to drafting and/or revising the manuscript.

Funding This work includes independent financial support by the Assistance Publique–Hôpitaux de Paris (AP-HP), Institut National de la Santé et de la Recherche Médicale (Inserm) and Sorbonne Université.

Competing interests Professor Devilliers reports personal fees from Johnson & Johson, personal fees from GSK, grants from LFB, non-financial support from Novartis, non-financial support from MSD, outside the submitted work

Ethics approval The study protocol was approved by the Comité de Protection des Personnes (CPP) EST I (Dijon). The study was conducted in accordance with the ethical principles of the Declaration of Helsinki.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

ORCID iD
Alexis Mathian http://orcid.org/0000-0002-7653-6528

REFERENCES

1 van Vollenhoven RF, Mosca M, Bertsias G, *et al*. Treat-to-target in systemic lupus erythematosus: recommendations from an international Task force. *Ann Rheum Dis* 2014;73:958–67.
2 Fanouriakis A, Kostopoulou M, Alunno A, *et al*. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. *Ann Rheum Dis* 2019;78:736–45.
3 Bertsias GK, Ioannidis JPA, Aringer M, *et al*. EULAR recommendations for the management of systemic lupus erythematosus with neuropsychiatric manifestations: report of a task force of the EULAR standing Committee for clinical Affairs. *Ann Rheum Dis* 2010;69:2074–82.
4 Bertsias GK, Tektonidou M, Amoura Z, *et al*. Joint European League against rheumatism and European renal Association–European dialysis and transplant association (EULAR/ERA-EDTA) recommendations for the management of adult and paediatric lupus nephritis. *Ann Rheum Dis* 2012;71:1771–82.
5 Ngamjanyaporn P, McCarthy EM, Sergeant JC, *et al*. Clinicians approaches to management of background treatment in patients with SLE in clinical remission: results of an international observational survey. *Lupus Sci Med* 2017;4:e000173.
6 Mosca M, Tani C, Carli L, *et al*. Glucocorticoids in systemic lupus erythematosus. *Clin Exp Rheumatol* 2011;29:S126–9.
7 Zonana-Nacach A, Barr SG, Magder LS, *et al*. Damage in systemic lupus erythematosus and its association with corticosteroids. *Arthritis Rheum* 2000;43:1801–8.
8 Thamer M, Hernán MA, Zhang Y, *et al*. Prednisone, lupus activity, and permanent organ damage. *J Rheumatol* 2009;36:560–4.
9 Al Sawah S, Zhang X, Zhu B, *et al*. Effect of corticosteroid use by dose on the risk of developing organ damage over time in systemic lupus erythematosus—the Hopkins lupus cohort. *Lupus Sci Med* 2015;2:e000066.
10 Apostolopoulos D, Kandane-Rathnayake R, Raghunath S, *et al*. Independent association of glucocorticoids with damage accrual in SLE. *Lupus Sci Med* 2016;3:e000157.
11 Tsang-A-Sjoe MWP, Bultink IEM, Heslinga M, *et al*. Both prolonged remission and lupus low disease activity state are associated with reduced damage accrual in systemic lupus erythematosus. *Rheumatology* 2017;56:121–8.
12 Hochberg MC. Updating the American College of rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1725;1997.
13 Buyon JP, Petri MA, Kim MY, *et al*. The effect of combined estrogen and progesterone hormone replacement therapy on disease activity in systemic lupus erythematosus: a randomized trial. *Ann Intern Med* 2005;142:953–62.
14 Petri M, Kim MY, Kalunian KC, *et al*. Combined oral contraceptives in women with systemic lupus erythematosus. *N Engl J Med* 2005;353:2550–8.
15 Isenberg DA, Rahman A, Allen E, *et al*. BILAG 2004. development and initial validation of an updated version of the British Isles lupus assessment group's disease activity index for patients with systemic lupus erythematosus. *Rheumatology* 2005;44:902–6.
16 Yee C-S, Farewell V, Isenberg DA, *et al*. The BILAG-2004 index is sensitive to change for assessment of SLE disease activity. *Rheumatology* 2009;48:691–5.
17 Steiman AJ, Urowitz MB, Ibañez D, *et al*. Prolonged clinical remission in patients with systemic lupus erythematosus. *J Rheumatol* 2014;41:1808–16.
18 Gordon C*et al*. Definition and treatment of lupus flares measured by the BILAG index. *Rheumatology* 2003;42:1372–9.
19 Isenberg DA, Allen E, Farewell V, *et al*. An assessment of disease flare in patients with systemic lupus erythematosus: a comparison of BILAG 2004 and the flare version of SELENA. *Ann Rheum Dis* 2011;70:54–9.
20 Gladman D, Ginzler E, Goldsmith C, *et al*. The development and initial validation of the systemic lupus international collaborating clinics/American College of rheumatology damage index for systemic lupus erythematosus. *Arthritis Rheum* 1996;39:363–9.
21 van Vollenhoven R, Voskuyl A, Bertsias G, *et al*. A framework for remission in SLE: consensus findings from a large international Task force on definitions of remission in SLE (DORIS). *Ann Rheum Dis* 2017;76:554–61.
22 Wilhelm TR, Magder LS, Petri M. Remission in systemic lupus erythematosus: durable remission is rare. *Ann Rheum Dis* 2017;76:547–53.
23 Zen M, Iaccarino L, Gatto M, *et al*. Prolonged remission in Caucasian patients with SLE: prevalence and outcomes. *Ann Rheum Dis* 2015;74:2117–22.
24 Galbraith L, Manns B, Hemmelgarn B, *et al*. The steroids in the maintenance of remission of proliferative lupus nephritis (SIMPL) pilot trial. *Can J Kidney Health Dis* 2014;1.
25 Urowitz MB, Feletar M, Bruce IN, *et al*. Prolonged remission in systemic lupus erythematosus. *J Rheumatol* 2005;32:1467–72.
26 Medina-Quiñones CV, Ramos-Merino L, Ruiz-Sada P, *et al*. Analysis of complete remission in systemic lupus erythematosus patients over a 32-year period. *Arthritis Care Res* 2016;68:981–7.
27 Mok CC, Ho LY, Tse SM, *et al*. Prevalence of remission and its effect on damage and quality of life in Chinese patients with systemic lupus erythematosus. *Ann Rheum Dis* 2017;76:1420–5.
28 Tani C, Elefante E, Signorini V, *et al*. Glucocorticoid withdrawal in systemic lupus erythematosus: are remission and low disease activity reliable starting points for stopping treatment? A real-life experience. *RMD Open* 2019;5:e000916.
29 Goswami RP, Sit H, Ghosh P, *et al*. Steroid-Free remission in lupus: myth or reality; an observational study from a tertiary referral centre. *Clin Rheumatol* 2019;38:1089–97.
30 Walsh M, Merkel PA, Mahr A, *et al*. Effects of duration of glucocorticoid therapy on relapse rate in antineutrophil cytoplasmic antibody-associated vasculitis: a meta-analysis. *Arthritis Care Res* 2010;62:1166–73.

31 Tengstrand B, Larsson E, Klareskog L, *et al*. Randomized withdrawal of long-term prednisolone treatment in rheumatoid arthritis: effects on inflammation and bone mineral density. *Scand J Rheumatol* 2007;36:351–8.
32 Pincus T, Swearingen CJ, Luta G, *et al*. Efficacy of prednisone 1-4 mg/day in patients with rheumatoid arthritis: a randomised, double-blind, placebo controlled withdrawal clinical trial. *Ann Rheum Dis* 2009;68:1715–20.
33 Ugarte-Gil MF, Acevedo-Vásquez E, Alarcón GS, *et al*. The number of flares patients experience impacts on damage accrual in systemic lupus erythematosus: data from a multiethnic Latin American cohort. *Ann Rheum Dis* 2015;74:1019–23.
34 Stoll T*et al*. Analysis of the relationship between disease activity and damage in patients with systemic lupus erythematosus—a 5-yr prospective study. *Rheumatology* 2004;43:1039–44.
35 Mok CC, Ho LY, Cheung MY, *et al*. Effect of disease activity and damage on quality of life in patients with systemic lupus erythematosus: a 2-year prospective study. *Scand J Rheumatol* 2009;38:121–7.10.1080/03009740802415527
36 Corneloup M, Maurier F, Wahl D, *et al*. Disease-specific quality of life following a flare in systemic lupus erythematosus: an item response theory analysis of the French equal cohort. *Rheumatology* 2019. doi:10.1093/rheumatology/kez451. [Epub ahead of print: 17 Oct 2019].
37 Gladman DD, Urowitz MB, Rahman P, *et al*. Accrual of organ damage over time in patients with systemic lupus erythematosus. *J Rheumatol* 2003;30:1955–9.
38 Ruiz-Arruza I, Ugarte A, Cabezas-Rodriguez I, *et al*. Glucocorticoids and irreversible damage in patients with systemic lupus erythematosus. *Rheumatology* 2014;53:1470–6.
39 Bruce IN, O'Keeffe AG, Farewell V, *et al*. Factors associated with damage accrual in patients with systemic lupus erythematosus: results from the systemic lupus international collaborating clinics (SLICC) inception cohort. *Ann Rheum Dis* 2015;74:1706–13.
40 Yee C-S, Su L, Toescu V, *et al*. Birmingham SLE cohort: outcomes of a large inception cohort followed for up to 21 years. *Rheumatology* 2015;54:836–43.
41 Lim LSH, Pullenayegum E, Lim L, *et al*. From childhood to adulthood: the trajectory of damage in patients with juvenile-onset systemic lupus erythematosus. *Arthritis Care Res* 2017;69:1627–35.
42 Chen H-L, Shen L-J, Hsu P-N, *et al*. Cumulative burden of glucocorticoid-related adverse events in patients with systemic lupus erythematosus: findings from a 12-year longitudinal study. *J Rheumatol* 2018;45:83–9.
43 Strehl C, Bijlsma JWJ, de Wit M, *et al*. Defining conditions where long-term glucocorticoid treatment has an acceptably low level of harm to facilitate implementation of existing recommendations: viewpoints from an EULAR Task force. *Ann Rheum Dis* 2016;75:952–7.
44 Burmester GR, Buttgereit F, Bernasconi C. Evaluating the safety and efficacy of blinded tapering versus continuation of long-term prednisone (5 mg/d) in patients with rheumatoid arthritis who achieved low disease activity or remission on tocilizumab. Abstract #OP0030: presented at the EULAR Annual meeting, Madrid, June 12-15; 2019.
45 Costedoat-Chalumeau N, Galicier L, Aumaître O, *et al*. Hydroxychloroquine in systemic lupus erythematosus: results of a French multicentre controlled trial (plus study). *Ann Rheum Dis* 2013;72:1786–92.

CLINICAL SCIENCE

IFN-α kinoid in systemic lupus erythematosus: results from a phase IIb, randomised, placebo-controlled study

Frederic A Houssiau ,[1,2] Aikaterini Thanou,[3] Minodora Mazur,[4] Edgar Ramiterre,[5] Danny Alexis Gomez Mora,[6] Maria Misterska-Skora,[7] Risto Alfredo Perich-Campos,[8,9] Svetlana A Smakotina,[10] Sergio Cerpa Cruz,[11] Bassem Louzir,[12] Thérèse Croughs,[13] Michael Lucas Tee[14]

Handling editor Josef S Smolen

► Additional material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-216379).

For numbered affiliations see end of article.

Correspondence to
Professor Frederic A Houssiau, Pôle de Pathologies Rhumatismales Inflammatoires et Systémiques, Institut de Recherche Expérimentale et Clinique, Université catholique de Louvain, Brussels, Belgium; frederic.houssiau@uclouvain.be

The results of the phase IIb in systemic lupus erythematosus were presented at the 13th International Congress on Systemic Lupus Erythematosus (LUPUS 2019), San Francisco, California, USA (abstract #198).

Received 3 October 2019
Revised 28 November 2019
Accepted 30 November 2019
Published Online First 23 December 2019

To cite: Houssiau FA, Thanou A, Mazur M, *et al*. *Ann Rheum Dis* 2020;**79**:347–355.

ABSTRACT

Objective To evaluate the efficacy and safety of the immunotherapeutic vaccine interferon-α kinoid (IFN-K) in a 36-week (W) phase IIb, randomised, double-blind, placebo (PBO)-controlled trial in adults with active systemic lupus erythematosus (SLE) despite standard of care.

Methods Patients with SLE (185) with moderate to severe disease activity and positive interferon (IFN) gene signature were randomised to receive IFN-K or PBO intramuscular injections (days 0, 7 and 28 and W12 and W24). Coprimary endpoints at W36 were neutralisation of IFN gene signature and the BILAG-Based Composite Lupus Assessment (BICLA) modified by mandatory corticosteroid (CS) tapering.

Results IFN-K induced neutralising anti-IFN-α2b serum antibodies in 91% of treated patients and reduced the IFN gene signature ($p<0.0001$). Modified BICLA responses at W36 did not statistically differ between IFN-K (41%) and PBO (34%). Trends on Systemic Lupus Erythematosus Responder Index-4, including steroid tapering at W36, favoured the IFN-K and became significant ($p<0.05$) in analyses restricted to patients who developed neutralising anti-IFN-α2b antibodies. Attainment of lupus low disease activity state (LLDAS) at W36 discriminated the two groups in favour of IFN-K (53% vs 30%, $p=0.0022$). A significant CS sparing effect of IFN-K was observed from W28 onwards, with a 24% prednisone daily dose reduction at W36 in IFN-K compared with PBO ($p=0.0097$). The safety profile of IFN-K was acceptable.

Conclusions IFN-K induced neutralising anti-IFN-α2b antibodies and significantly reduced the IFN gene signature with an acceptable safety profile. Although the clinical coprimary endpoint was not met, relevant secondary endpoints were achieved in the IFN-K group, including attainment of LLDAS and steroid tapering.

Trial registration number NCT02665364.

Key messages

What is already known about this subject?
- Type I interferons (IFNs) play a pathogenic role in systemic lupus erythematosus (SLE).
- Interferon-α kinoid (IFN-K), an immunotherapeutic agent, elicits the production of anti-IFN-α antibodies.

What does this study add?
- IFN-K induced a strong polyclonal immunogenic response directed against IFN-α in nearly all patients and significantly reduced the IFN gene signature in the IFN-K group compared with placebo (PBO).
- The safety profile was acceptable.

How might this impact on clinical practice orfuture developments?
- In a 36-week phase IIb trial performed in patients suffering from active SLE, treatment with IFN-K did not increase the percentage of BILAG-Based Composite Lupus Assessment responders (clinical coprimary endpoint) but allowed more steroid reduction. Lupus low disease activity state was achieved in more patients on IFN-K compared with PBO.

INTRODUCTION

The pivotal role of type I interferons (IFNs) in pathogenesis of systemic lupus erythematosus (SLE) has been the focus of extensive research spanning two decades.[1–3] Despite promising preclinical evidence, results of clinical trials of several type I IFN blockers in SLE have been mixed.[4–7] Rontalizumab, a monoclonal antibody against IFN-α, did not meet its primary endpoint (Systemic Lupus Erythematosus Responder Index-4 (SRI-4)) in a phase II trial, although exploratory analyses indicated clinical benefit and steroid-sparing effects in the subset of patients with a lower IFN gene signature at baseline, reflecting IFN-regulated gene expression.[5] Sifalimumab, a fully human monoclonal antibody against most IFN-α subtypes, achieved its primary endpoint (SRI-4) in a phase IIb study, but differences were only modest.[6] Anifrolumab, a monoclonal antibody against the type I IFN receptor that inhibits signalling of all type I IFNs, was superior to placebo (PBO) across multiple endpoints in a phase IIb trial.[7] One of the two phase III trials (TULIP-1), using SRI-4 as primary outcome measure, did not confirm these results,[8] while the other (TULIP-2)[9] achieved an alternative primary endpoint, the BILAG-Based Composite Lupus Assessment (BICLA). Baricitinib

inhibits Janus kinase 1/2 affecting multiple cytokines but also downstream signalling through type I IFNs, and was found to be superior to PBO for SLE arthritis and rash in a recent phase II trial.[10]

The interferon-α kinoid (IFN-K) is an immunotherapeutic vaccine composed of inactivated recombinant human IFN-α2b coupled to a T-helper carrier protein (keyhole limpet haemocyanin), aimed at inducing antibodies against IFN-α by active immunisation. When injected intramuscularly in human IFN-α transgenic mice, IFN-K yielded a strong polyclonal response, targeting multiple epitopes, enabling to recognise not only IFN-α2b but also the 12 other human IFN-α subtypes.[11] Accordingly, IFN-K was shown to slow disease progression in a mouse model of SLE.[12] In a phase I/IIa dose-escalation PBO-controlled study in patients with active SLE, IFN-K was well tolerated, induced high titres of neutralising anti-IFN-α antibodies, especially in patients with a type I IFN signature, and significantly reduced expression of IFN-induced genes.[13] Follow-up analyses on a subgroup of IFN-K-treated patients confirmed the link between persistence of anti-IFN-α antibodies and down-regulation of the IFN signature and revealed an inhibitory effect of IFN blockade on B cell-associated transcripts.[14]

Here, we present and discuss the results of a 36-week (W) phase IIb, randomised, double-blind, PBO-controlled, multicenter study, designed to assess efficacy and safety of IFN-K in patients with active SLE despite standard of care.

PATIENTS AND METHODS

Study design

This study was a W36 randomised, double-blind, PBO-controlled (1:1), multinational (22 countries), phase IIb trial evaluating the neutralisation of the IFN gene signature and the clinical efficacy of IFN-K in adults with SLE. The protocol was approved by an independent institutional review board at each participating site, and all patients signed informed consent before any study-related procedures. An independent data safety monitoring board, consisting of experts in the appropriate disciplines, oversaw patient's safety every 6 months and ad hoc in case of emerging safety concerns. An adjudication committee of independent experts confirmed the accuracy and consistency of the British Isles Lupus Assessment Group (BILAG)−2004 Index.[15]

Inclusion/exclusion criteria

All of the following inclusion criteria were required: age 18–65 years; SLE of ≥4/11 by 1997 American College of Rheumatology criteria,[16] Systemic Lupus Erythematosus Disease Activity Index-2000 (SLEDAI-2K)≥6[17]; ≥1 BILAG A and/or ≥2 BILAG B scores; positive IFN gene signature antinuclear antibody titre of ≥1:160 and/or positivity of antidouble-stranded deoxyribonucleic acid (dsDNA) antibodies; treatment with at least one of the following: corticosteroids (CSs) at ≤20 mg of prednisone equivalent/day, hydroxychloroquine (HCQ) or chloroquine on stable dose for at least 4W prior to the first planned study drug administration, mycophenolate mofetil (MMF)/mycophenolic acid (≤2 g/day), methotrexate (≤20 mg/week) and azathioprine (≤2.5 mg/kg/day), all on stable dose for at least 12W prior to the first study drug administration. The following exclusion criteria were applied: active severe lupus nephritis (renal BILAG A or immediate need for cyclophosphamide), active severe neuropsychiatric lupus, treatment with >20 mg of prednisone equivalent/day for >7 consecutive days within 4 months prior to the first study drug administration, pulse CS (≥250 mg prednisone/day) within 3 months prior to the first study drug administration, treatment with cyclophosphamide, cyclosporine A, tacrolimus, abatacept, sifalimumab, rontalizumab, anifrolumab, belimumab, tumour necrosis factor antagonists, anti-B cell therapy or any other registered investigational biological therapy or live vaccine, and use of investigational non-registered product or investigational non-registered vaccine within 3–12 months (according to drug) prior to the first study drug administration. Additional exclusion criteria included ≥6 occurrences of oral/genital herpes simplex virus (HSV) infections or any episode of shingles within 12 months prior to the first study drug administration, absence of IgG anti-HSV 1/2, antivaricella zoster virus (VZV), anticytomegalovirus (CMV) and anti-Epstein-Barr virus (EBV) serum antibodies at screening, presence of anti-HTLV1/2, anti-HIV, antihepatitis C virus (HCV) serum antibodies or hepatitis B surface antigen at screening, anticipated high risk of significant infection by physician's opinion, current signs or symptoms of infection, treatment with intravenous antibiotics within 2 months prior to the first planned study drug administration, high-risk human papilloma virus (HPV) positivity on a cervical swab by real-ime quantitative polymerase chain reaction (RT-qPCR) and cytological abnormalities of ≥high-grade superficial intraepithelial lesion. Fibromyalgia was not an exclusion criterion but was reported at study entry in only five patients (three and two patients in the IFN-K and PBO groups, respectively).

Treatment

After a 4W screening period, patients with SLE were randomised to the IFN-K or PBO group using a minimisation algorithm by age, ethnicity, presence of renal involvement and treatment with CS and/or HCQ and/or MMF. They received five intramuscular injections of IFN-K or an equivalent volume of 0.9% NaCl, both emulsified with an oil-based adjuvant (Montanide, ISA 51VG, Seppic, France): 240 µg at days (D) 0, 7 and 28 and 120 µg at W12 and W24. CS administration was strictly controlled, with a maximum dose of 20 mg equivalent prednisone/day at D0, a recommended taper to 10 mg/day by W12 and a mandatory target of ≤5 mg/day by W24 without further increase until W36. Patients not fulfilling the CS tapering rule between W24 and W36 remained in the study but were considered treatment failures.

Efficacy and safety evaluations

Coprimary efficacy measures at W36 compared with baseline were (1) neutralisation of IFN gene signature, measured by change in the expression of IFN-induced genes, and (2) clinical response measured by the BICLA[18] with superimposed CS tapering (≤5 mg prednisone equivalent/day at W24 with no increase until W36, modified BICLA). Secondary efficacy measures at W36 were SRI-4,[19] SRI-4 with CS tapering (≤5 or ≤7.5 mg/day), SLEDAI-2K, Safety of Oestrogen in Lupus Erythematosus National Assessment (SELENA)–Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) flare index, BILAG-2004 Index, Systemic Lupus International Collaborating Clinics/ACR Damage Index for SLE,[20] Cutaneous Lupus Erythematosus Disease Area and Severity Index (CLASI)[21] in patients with cutaneous lesions at baseline, lupus low disease activity state (LLDAS)[22] and health-related quality of life assessed by SF36.[23] Safety was evaluated by the incidence, nature, severity and drug relatedness of adverse events (AEs).

IFN gene signature

At screening, from blood samples, IFN gene signature was tested using RT-qPCR, on a selection of 10 IFN-inducible genes

(*IFIT3*, *MX1*, *ISG15*, *IFIT1*, *IFI6*, *OAS2*, *HERC5*, *LY6E*, *IFI27* and *SIGLEC1*) known to strongly correlate with the IFN signature based on the 21-probe set described by Yao *et al*.[24] Positive signature was defined by a fold change of ≥3 compared with healthy donor blood samples. At randomisation (D0) and W12, W24 and W36, the IFN gene signature, including the 21 IFN-regulated genes, was evaluated by Affymetrix.[24] RNA was extracted, quality assessed, labelled, hybridised to GeneChip human genome U133 Plus 2.0 arrays, washed, stained and scanned using GeneChip Scanner 3000 7G. Affymetrix CEL files were uploaded to Affymetrix Expression Console software V.1.4.1. Analysis and Robust multichip Analysis normalisation of raw data were performed per batch, and raw dataset was normalised at once. Data are accessible on request.

Serum anti-IFN-α antibodies and anti-IFN-α neutralising capacity

Serum anti-IFN-α2b antibody titres were measured by ELISA as described elsewhere,[11–14] every 4W from D0 onwards. Serum antibody neutralising capacity against recombinant IFN-α2b and 12 other IFN-α subtypes was measured by reporter gene assay using interferon-sensitive response element reporter, HEK293 cells containing the firefly luciferase gene. Neutralising capacity corresponds to the first dilution factor of sera resulting in 50% neutralisation of IFN-induced luminescence (30 U/mL). Results were expressed as the highest dilution of the serum in which antibodies could be detected. The lowest dilution tested (and the limit for positivity) was 1/400 and 1/200 for binding and neutralising antibodies, respectively.

Statistical analyses

Efficacy and safety analyses were performed on patients who received at least one dose of the study drug. The study was considered positive if there was superiority of IFN-K in neutralising the IFN gene signature and a ≥10% difference favouring the IFN-K on the modified BICLA response. A sample size of 160 evaluable patients (80 patients per group) would provide a 85% power at detecting a 32.6% difference of IFN-K over PBO in IFN-induced gene expression, using a two-group t-test at a 0.05 two-sided significance level, assuming a common SD of 68%. Assuming a 40.6% BICLA response on IFN-K and a 20.6% response on PBO, that sample size would also provide a 73% power at detecting a 20% BICLA response difference between groups. The biological coprimary endpoint was analysed using a covariance model, with percentage change from baseline in the expression of IFN-induced genes as dependent variable and treatment assignment as independent variable. Minimisation factors used for randomisation were included as covariates. Modified BICLA and SRI-4 were analysed by logistic regression, with response as dependent variable and treatment assignment as independent variable, adjusting for minimisation factors. SLEDAI response (≥4-point reduction in SLEDAI-2K) at W36 versus baseline was compared between groups by frequency table methods. LLDAS at W36 in each group was assessed by Pearson χ^2 test. Anti-IFN-α antibody titres and their neutralising capacity over time were analysed using frequency table methods. Statistical analyses were performed by SAS® software V.9.4.

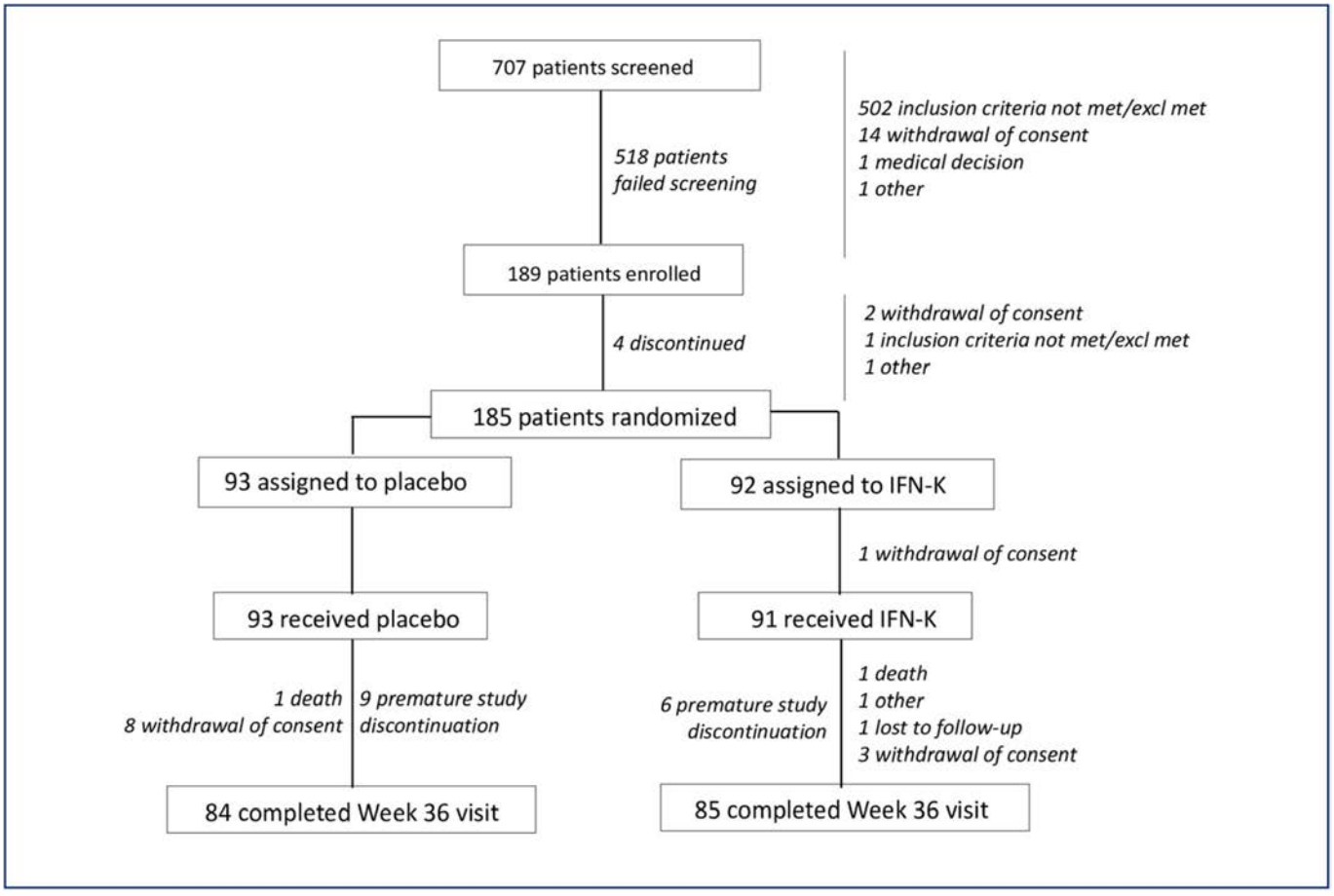


Figure 1 Patients' disposition. Screen failures are detailed in online supplementary material 1. IFN, interferon.

Table 1 Demographics and baseline characteristics

	IFN-K (n=91)	PBO (n=93)
Age (years)	39.53±10.30	38.75±11.16
Gender		
Male, n (%)	7 (7.7)	5 (5.4)
Female, n (%)	84 (92.3)	88 (94.6)
Ethnicity		
Black, n (%)	1 (1.1)	1 (1.1)
Asian, n (%)	16 (17.6)	10 (10.8)
Caucasian/Hispanic, n (%)	64 (70.3)	66 (71.0)
Other, n (%)	10 (11.0)	16 (17.2)
Time since diagnosis (years)	6.7±6.4	7.1±6.6
IFN gene signature (score) at baseline	4.5±1.0	4.5±1.3
SLEDAI-2K Global Score	10.3±3.7	11.3±4.0
BILAG-2004 Index		
BILAG-2004 Index Global Score	18.2±8.2	18.5±5.9
Mucocutaneous BILAG A, n (%)	14 (15.4)	12 (12.9)
Mucocutaneous BILAG B, n (%)	67 (73.6)	66 (71.0)
Musculoskeletal BILAG A, n (%)	17 (18.7)	20 (21.5)
Musculoskeletal BILAG B, n (%)	56 (61.5)	59 (63.4)
Physician's global assessment (mm)	56.3±17.7	53.1±18.1
CLASI Total Activity Score	5.4±6.3	5.3±4.6
Joint pain VAS (mm)	44.8±23.3	46.5±21.9
28-Tender joints count, n	7.3±6.4	7.9±6.5
28-Swollen joints count, n	4.8±4.7	4.6±3.6
FACIT Fatigue Score	32.7±10.0	32.5±10.8
Complement C3 (mg/L)	872±271	909±297
Complement C4 (mg/L)	135±70	151±90
Anti-dsDNA (Phadia, U/mL)	456±2367	101±264
Concomitant medications at scrrening		
Corticosteroids, n (%)	82 (90.1)	84 (90.3)
Mean daily prednisone dose (mg)	8.9	9.3
Hydroxychloroquine, n (%)	60 (65.9)	69 (74.2)
Mycophenolate mofetil, n (%)	9 (9.9)	7 (7.5)
Methotrexate, n (%)	11 (12.1)	14 (15.1)
Azathioprine, n (%)	15 (16.5)	16 (17.2)

Unless stated otherwise, data are means ± SD. No statistical differences were observed between the two treatment groups.
BILAG, British Isles Lupus Assessment Group; CLASI, Cutaneous Lupus Erythematosus Disease Area and Severity Index; FACIT, Functional Assessment of Chronic Illness Therapy; IFN, interferon; IFN-K, interferon-α kinoid; PBO, placebo; SLEDAI-2K, Systemic Lupus Erythematosus Disease Activity Index-2000; VAS, Visual Analogue Scale.

RESULTS

Study population

As depicted in figure 1, 185 patients with active SLE with a positive IFN gene signature were randomised, with 93 assigned to PBO and 92 to IFN-K. Reasons for screen failures are detailed in online supplementary material 1. As expected, one-third of patients were excluded because of the absence of IFN signature. Absence of serum IgG antibodies against HSV, VZV, CMV or EBV and/or detection by RT-qPCR of high-risk HPV on a cervical swab also contributed to the high rate of screen failure. Patient demographic and baseline clinical characteristics are described in table 1. The mean age was 39 years. Majority of the patients were female (94%) and Caucasian (71%). Most individuals suffered from mucocutaneous and musculoskeletal disease (BILAG A or B) despite standard of care, including CS, HCQ and/or other immunosuppressants in 90%, 70% and 39% of them, respectively.

Immunogenicity

In the IFN-K group, 98% of the patients developed anti-IFN-α2b-binding antibodies with a titre ≥1/400 at W36. Among them, 50% patients had a serum titre above 1/25 600, 28% above 1/51 200 and 10% above 1/102 400. Neutralising anti-IFN-α2b antibodies were detected in 71% of patients (serum titre ≥1/200) as early as W12, and these were detected in 91% of IFN-K-treated patients at W36 (figure 2A), with 50% displaying a titre above 1/12 800, 30% above 1/25 600, 14% above 1/51 200 and 5% above 1/102 400. The IFN-K also induced polyclonal antibodies able to neutralise other IFN-α subtypes in 30%–97% of treated patients (figure 2B). Taken together, more than half of the patients treated with IFN-K raised a neutralising response against 9–12 IFN-α subtypes. No cross neutralisation against IFN-β was observed, and only some weak cross reactivity was detected against IFN-ω in five patients at W36, with one patient titre at 1/100, two patients at 1/200 and two patients at 1/800.

Efficacy

Treatment with IFN-K induced a statistically significant 31% mean reduction from baseline in the expression of type I IFN gene score at W36 (figure 3A), which was not observed in PBO-treated patients ($p<0.0001$). Of note, 20/87 patients did experience an increase of IFN gene signature (mean +23%), which is likely related to a lower immune response against IFN-α. Patients with anti-IFN-α2b neutralising antibody titres between 1/100 and 1/1600 had indeed a lower decrease in their signature (mean −8.23%) compared with patients with titres of >1/1600 (mean −42.4%). While the biological coprimary endpoint was met, the modified BICLA response difference in favour of IFN-K over PBO was only of 6.7% (figure 3B).

SRI-4 response at W36 also did not differ between treatment groups. Nevertheless, when CS restrictions were added, a trend favouring the IFN-K was observed (figure 4A). Combined with a requirement for CS tapering to ≤5 or ≤7.5 mg of prednisone equivalent/day (by W24, with no increase to W36), SRI-4 at W36 yielded a 15.4% ($p=0.076$) and a 15.3% ($p=0.079$) difference of IFN-K over PBO, respectively. This became significant in exploratory analyses restricted to patients with neutralising anti-IFN-α2b antibodies, with a 16.6% ($p=0.042$) and 16.8% difference ($p=0.0396$) respectively (figure 4A). At W36, 52.9% of patients assigned to IFN-K achieved LLDAS, which was reached in only 29.8% of PBO-treated patients. This 23% difference was highly significant ($p=0.0022$). Consistent with differences in composite endpoints when a CS target was included, the mean daily prednisone dose was significantly lower in IFN-K-treated patients from W28 onwards, with a 24% dose reduction from baseline at W36 ($p=0.009$) (figure 4B).

Evolution over time of SLEDAI-2K, BILAG-2004 Global Index, BILAG-2004 musculoskeletal and mucocutaneous domains, Physician Global Assessment (PGA), CLASI, tender joint count, swollen joint count, Complement C3 and C4, and anti-dsDNA are detailed in online supplementary materials 2 and 3. A trend favouring the IFN-K over PBO was observed in change in PGA (baseline to W36, $p=0.0537$). None of these measures, however, discriminated the two treatment groups. Mild and moderate disease flares, defined per SELENA–SLEDAI flare index, were observed in 9% and 12% of the IFN-K and PBO-treated patients, while severe disease flares were observed in 3% and 6% of IFN-K and PBO patients, respectively (data not shown). Achievement of clinical meaningful improvement in quality of life by 36-Item Short-Form Health Survey (SF-36) questionnaire (≥2.5 change from baseline in physical or mental

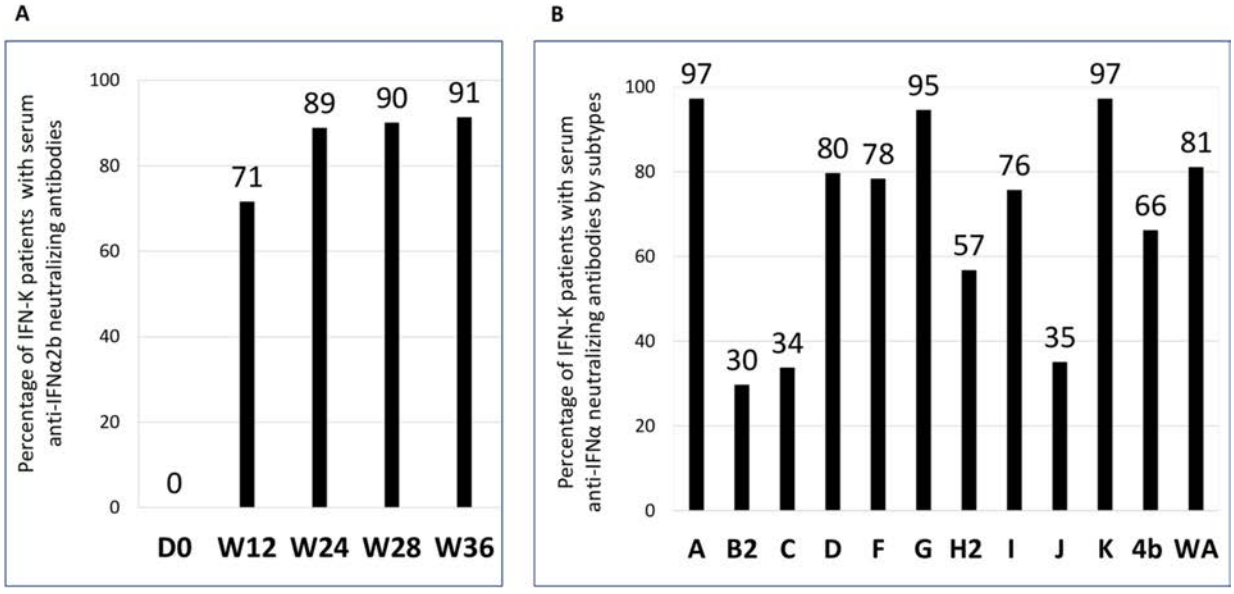


Figure 2 Induction of serum anti-IFN-α neutralising antibodies in IFN-K-treated patients. Percentages of IFN-K-treated patients with serum neutralising antibodies against IFN-α2b (A) and 12 other IFN-α subtypes (B) are indicated on top of each column. Kinetics is shown in (A), while W36 data are shown in (B). Serum titres (dilutions) of ≥1/400 (for binding antibodies) and ≥1/200 (for neutralising capacity) were considered positive. IFN, interferon; IFN-K, interferon-α kinoid.

component summaries or ≥5 for each subdomain) did not differ between groups (online supplementary material 4), except for a trend (p=0.068) favouring IFN-K in the energy/fatigue domain.

Safety and tolerability

Adverse events (AE), treatment-emergent adverse events (TEAEs), severe TEAEs, TEAEs leading to permanent study drug discontinuation and related serious adverse events (SAEs) were equally distributed between the two groups, as indicated in table 2. Related AEs were more frequent with IFN-K (40.7%) than with PBO (24.7%). SAEs were more frequent with PBO (12.9%) than with IFN-K (6.6%). Treatment-emergent serious adverse event (TESAE) leading to permanent study drug discontinuation or of severe intensity were more frequent in PBO (3.2% and 6.5%) than IFN-K (1.1% and 3.3%), respectively. Two patients died, one from pneumonia and lupus disease progression (IFN-K group) and one from central nervous system lymphoma (PBO group). Four cases of cancer were observed in the PBO group and none in the IFN-K group. Shingles were observed in two patients on IFN-K and one on PBO. One patient on IFN-K experienced a severe episode of rash referred to as a

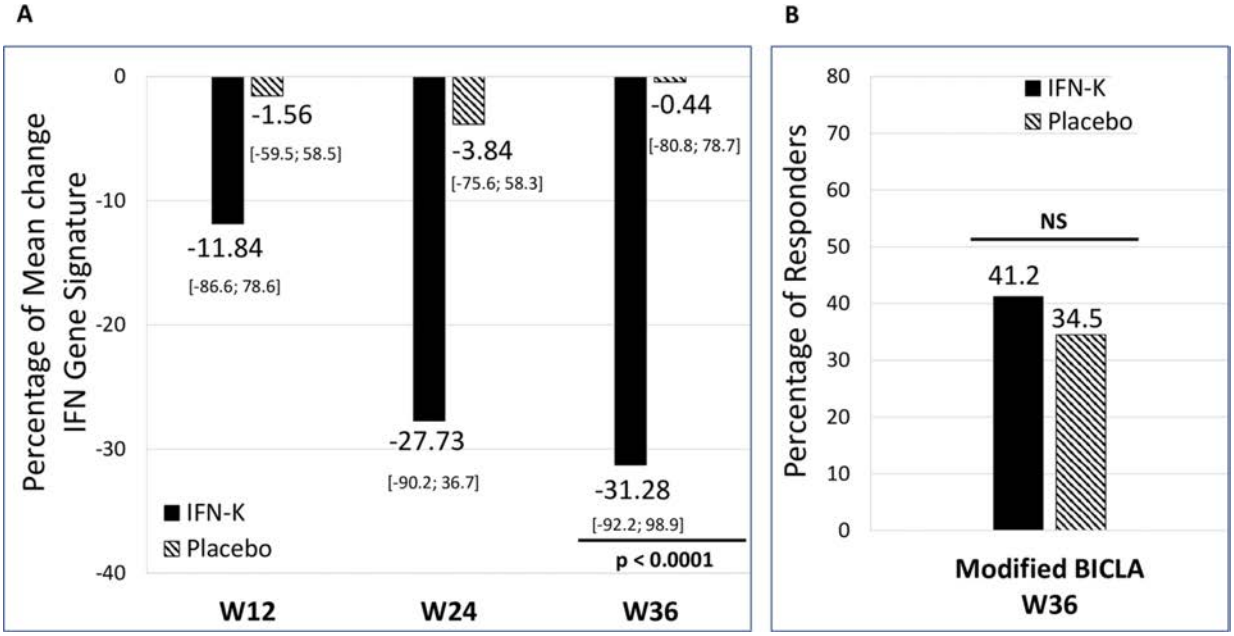


Figure 3 Coprimary endpoints at W36. Mean (min and max) percentages of change from baseline expression of IFN-induced genes (biological coprimary endpoint), evaluated by Affymetrix at W12, W24 and W36, in patients treated with IFN-K (closed columns) and placebo (hatched columns) are shown in (A). P<0.0001 by ANCOVA model (primary readout). Percentages of modified BICLA responders at W36 (clinical coprimary endpoint) are illustrated in (B). As explained in the Patients and methods section, BICLA was modified by the addition of a corticosteroid tapering rule, namely, a ≤5 mg/day prednisone target dose at W24, without further increase until W36. BICLA, BILAG-Based Composite Lupus Assessment; IFN, interferon; IFN-K, interferon-α kinoid.

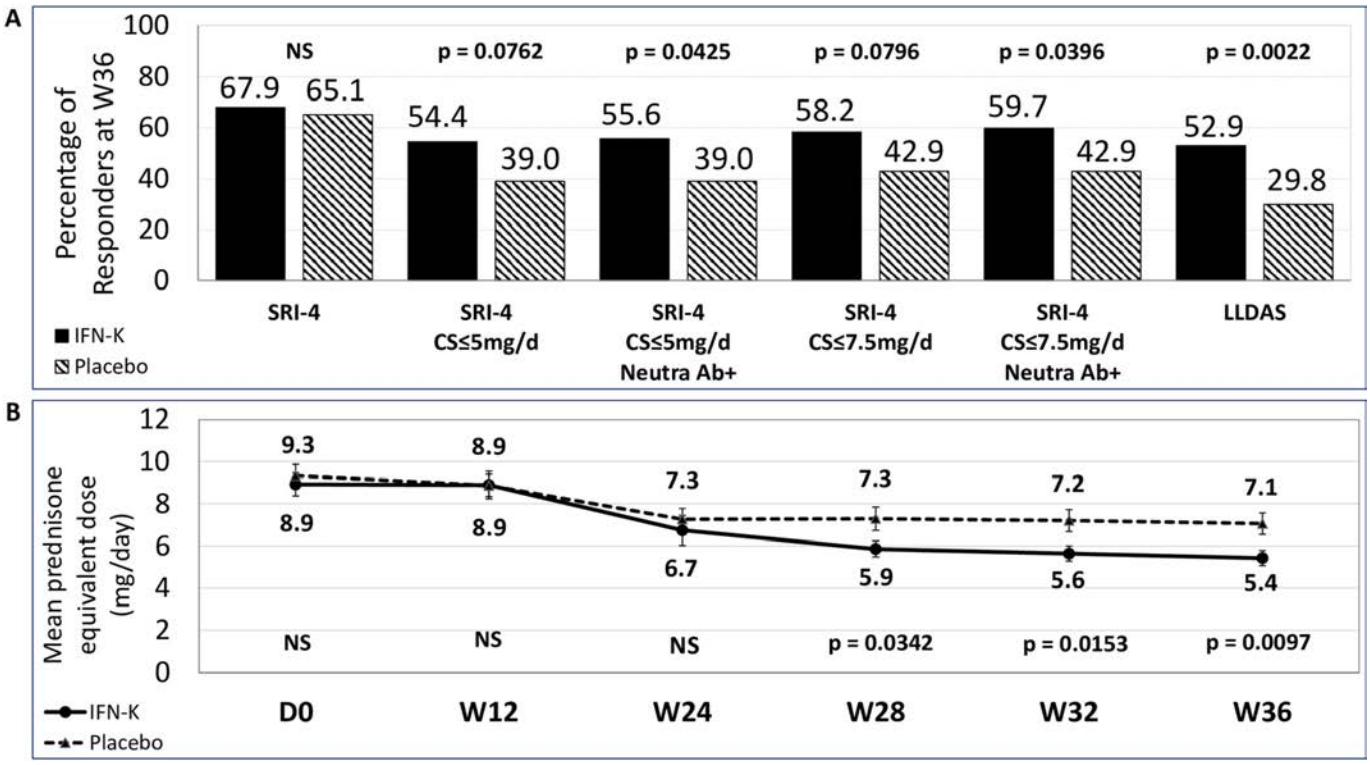


Figure 4 Main secondary endpoints and exploratory analyses at W36. Percentages of SRI-4 and LLDAS responders at W36 in IFN-K-treated patients (closed columns) and those in PBO-treated patients (hatched columns) are shown in (A). As indicated, a trend in favour of IFN-K was observed when a CS target was added to the SRI-4 endpoint (logistic regression model Wald χ^2), which became significant in exploratory analyses when the five patients who did not raise serum neutralising anti-IFN-α2b antibodies were excluded from the analyses (Pearson χ^2). Mean daily prednisone equivalent doses over time in IFN-K-treated patients (continuous line) and PBO-treated patients (dotted line) are shown in (B). They statistically differ from W28 onwards: p=0.0342, 0.0153 and 0.0097 at W28, W32 and W36 respectively (Student-Satterthwaite). CS, corticosteroid; IFN, interferon; IFN-K, interferon-α kinoid; LLDAS, lupus low disease activity state; NS, PBO, placebo;SRI, Systemic Lupus Erythematosus Responder Index.

Kaposi varicelliform eruption, with full recovery except for cutaneous scars. Among TEAEs reported with >5% frequency in the IFN-K group, upper respiratory tract infections and arthralgia were three times more common in the IFN-K group, and nasopharyngitis was twice more common. Injection site induration was observed in 5.5% of IFN-K-treated patients.

DISCUSSION

In this phase IIb trial, the IFN-K induced neutralising anti-IFN-α serum antibodies and significantly down-regulated the IFN gene signature, achieving the biological coprimary endpoint. The clinical coprimary endpoint, that is, the modified BICLA response at W36, was not met. Nonetheless, secondary composite endpoints that incorporated a CS tapering rule favored the IFN-K group. This was observed for SRI-4 with CS tapering to ≤5 or ≤7.5 mg prednisone equivalent/day (by W24 with no increase to W36) and became significant in the subgroup with neutralising antibodies to IFN-K. Similarly, attainment of LLDAS, which also includes a requirement for CS tapering to ≤7.5 mg/day, was significantly in favour of the IFN-K at W36. This is important since LLDAS has been associated with reduced organ damage accrual,[22 25 26] improved quality of life[27] and reduced healthcare costs in SLE.[28]

A statistically significant and clinically relevant CS sparing effect was observed in the IFN-K-group from W28 onwards. Overall, the facilitation of CS tapering by the IFN-K treatment, while maintaining clinical efficacy, was a striking observation in this trial. Damage accrual in SLE has been linked to cumulative CS exposure,[29] and reducing the CS burden remains a major objective of patients themselves due to its adverse effects on their body image and self-esteem. CS taper has been therefore included in treat-to-target recommendations advocated by an international task force,[30] as well as the 2019 update of the EULAR recommendations for management of SLE.[31] While the underlying mechanisms pertaining to its CS sparing effect can only be speculated, it is plausible that, by blocking IFN-α and subsequently decreasing the expression of proteins involved in autoimmunity, IFN-K down-regulates disease activity, thereby allowing lowering of CS, the more so in the setting of a clinical trial where CS tapering is mandatory. In other words, the effects of IFN-K are unmasked by imposing CS reduction. Another hypothesis is that type I IFNs exert yet unknown inhibitory effects on CS-induced pathways. IFN-α inhibition by the IFN-K immunisation may therefore lead to enhanced CS efficacy, thereby allowing CS dose reduction. Evidence supporting this possibility includes the observation that Toll-like receptor (TLR)-induced activation of type I IFN pathways may be intrinsically CS-insensitive. It was indeed shown that IFN-induced genes are not suppressed by CS other than intravenous pulse doses because of TLR-activated NFκB being CS-resistant.[32] This hypothesis requires further experimental evaluation.

Previous studies of other type I IFN targeting therapies did not reveal overt and unexpected toxicities. Similarly, the safety profile of the IFN-K was quite acceptable in this study, with even less SAEs compared with PBO. Of note, HSV-seronegative, VZV-seronegative, CMV-seronegative and EBV-seronegative (IgG) patients were excluded from this trial as a cautionary measure against primary infections when all subsets of IFN-α could have

Table 2 AEs

	IFN-K (n=91)	PBO (n=93)
Any AE	78 (85.7%) (392)	74 (79.6%) (302)
Any TEAE	75 (82.4%) (371)	71 (76.3%) (277)
TEAE leading to study treatment permanent discontinuation	4 (4.4%) (4)	4 (4.3%) (4)
TEAE of intensity severe or more	10 (11.0%) (27)	10 (10.8%) (11)
Related TEAE	37 (40.7%) (95)	23 (24.7%) (54)
Any SAE	6 (6.6%) (13)	12 (12.9%) (15)
Any TESAE	6 (6.6%) (13)	12 (12.9%)(15)
TESAE leading to study treatment permanent discontinuation	1 (1.1%) (1)	3 (3.2%) (3)
TESAE of intensity severe or more	3 (3.3%) (9)	6 (6.5%) (6)
Related TESAE	2 (2.2%) (7)	2 (2.2%) (2)
Death	1 (1.1%) (2)	1 (1.1%) (1)
Adverse events of interest		
Herpes zoster	2 (2.2%) (2)	1 (1.1%) (1)
Severe Infection	2 (2.2%) (2)	0 (0.0%) (0)
Malignancy	0 (0.0%) (0)	4 (4.3%) (4)*
Most common adverse events (>5% in the IFN-K group) by PT		
Upper respiratory tract infection	16 (17.6%)(17)	5 (5.4%)(6)
Urinary tract infection	11 (12.1%)(11)	9 (9.7%)(10)
Nasopharyngitis	7 (7.7%)(10)	2 (2.2%)(2)
Pharyngitis	6 (6.6%)(7)	3 (3.2%)(4)
Bronchitis	5 (5.5%)(5)	4 (4.3%)(4)
Injection site induration	5 (5.5%)(8)	0 (0.0%)(0)
Arthralgia	7 (7.7%)(8)	2 (2.2%)(3)
Pain in extremity	6 (6.6%)(6)	1 (1.1%)(1)
Headache	10 (11.0%)(19)	2 (2.2%)(3)

Data are numbers (and percentage) of patients. Data in brackets () are numbers of events. More than 1 event can be reported per patient. No statistical differences were observed between the two groups.
AEs were considered as treatment emergent if date of event was at or after the date of the first study drug administration.
*Two papillary thyroid cancers, one central nervous system lymphoma, one rectal cancer.
AE, adverse event; IFN-K, interferon-α kinoid; PBO, placebo; PT, preferred term; SAE, serious adverse event; TEAE, treatment-emergent adverse event; TESAE, treatment-emergent serious adverse events.

been blocked. Bearing this limitation, the viral infection profile of the IFN-K was reassuring with no increased risk of viral infections compared with PBO. The ongoing follow-up study on patients who received IFN-K will inform us on the long-term safety and efficacy, as well as the kinetics of the IFN-K-induced anti-IFN-α response that is expected to variably wane over time.

The concept of blocking IFN-α by IFN-K is consistent with the paradigm of personalised medicine since we were able to demonstrate that patients with the strongest type I IFN signature at baseline mounted the strongest anti-IFN-α response.[13] Yet, this should not disguise the following limitations. First, only two-thirds of patients with SLE display a type I IFN signature, making them eligible for IFN-K therapy. Second, IFN-K does not block other IFN subtypes like IFN-ω and IFN-β or type-specific like type II (IFN-γ) or type III (IFN-λ), which may explain the level of the effect observed in this trial compared with other IFNs targeted therapies. Third, the kinetics of the persistence of blocking IFN-α antibodies clearly needs to be addressed, as well as the duration of the inhibition of the IFN signature. Finally, the lack of improvement of patient-reported outcomes, shared by other anti-IFN drugs, is puzzling and disappointing.

In summary, based on preclinical data obtained in murine lupus models, on translational research performed in patients with lupus and on clinical trials, type I IFNs and related pathways remain key targets for the treatment of active SLE. Indeed, of all molecules tested so far, rontalizumab,[5] sifalimumab,[6] anifrolumab[7] and baricitinib[10] have demonstrated some efficacy over PBO on one or more outcome measures (primary and/or secondary endpoints, total and/or subset population). Yet, as of today, none of these compounds have yielded positive results in more than one phase III studies, which is required for approval by medical drug agencies. It has been increasingly acknowledged that these failures may be more related to the choice of the outcome measures than to actual inefficacy of the molecules. The IFN-K study reported here further fuels this hypothesis, since the drug did not meet its primary endpoint despite a significant steroid-sparing effect and attainment of LLDAS, indicating that the IFN-K deserves further evaluation in phase III studies.

Author affiliations
[1]Pôle de Pathologies Rhumatismales Inflammatoires et Systémiques, Institut de Recherche Expérimentale et Clinique, Université Catholique de Louvain, Brussels, Belgium
[2]Service de Rhumatologie, Cliniques Universitaires Saint-Luc, Brussels, Belgium
[3]Oklahoma Medical Research Foundation Arthritis and Clinical Immunology Research Program, Oklahoma, Oklahoma, USA
[4]Department of Internal Medicine, State Medical and Pharmaceutical University 'Nicolae Testemitanu', Chisinau, Republic of Moldova
[5]Section of Rheumatology, Department of Internal Medicine, Southern Philippines Medical Center, Davao, Philippines
[6]Healthy Medical Center, Universidad Nacional de Colombia, Zipaquira, Colombia
[7]Medical Center Oporów, Wrocław, Poland
[8]Rheumatology Department, Hospital Nacional Guillermo Almenara Irigoyen, Lima, Peru
[9]Universidad Nacional Mayor de San Marcos, Lima, Peru
[10]Therapy Department of Kemerovo, Kemerovo State Medical University, Kemerovo, Russian Federation
[11]Rheumatology Department, Hospital Civil de Guadalajara 'Fray Antonio Alcalde', Guadalajara, Mexico
[12]Department of Internal Medicine, Military hospital of Tunis, Tunis, Tunisia
[13]Neovacs S.A, Paris, France
[14]Department of Physiology, College of Medicine, University of the Philippines Manila and ManilaMed, Manila, Philippines

Acknowledgements The Steering Committee members* and the authorsΔ sincerely thank the patients and all investigators and their site personnel who actively contributed to the study. Argentina: Leonardo Naftal; Belgium: Frédéric A Houssiau*Δ; Chile: Elizabeth Jean Moreno Goio, Lilian Andrea Soto Sáez, Marta Ofelia Aliste Silva; Colombia: Elias Gonzalo Forero Illera, Patricia Julieta Vélez Sánchez, Danny Alexis Gómez MoraΔ, Diego Luis Saaibi Solano; Croatia: Marinko Artuković; France: Laurent Chiche; Georgia: Nino Jeiranashvili, Nino Kiparoidze, Mamuka Lortkipanidze; Germany: Rieke Alten, Herbert Kellner; Italy: Lorenzo Cavagna; Fabrizio Conti, Nazzarena Malavolta, Marta Mosca; Mexico: Hilario Ernesto Avila Armengol, Ignacio García de la Torre, Sandra Miriam Carrillo Vázquez, Sergio Cerpa CruzΔ, Dalila Alejandra Lopez Rodriguez, José Antonio Veloz Aranda, Jorge Enrique Aguilar Arreola; Moldova: Mazur MinodoraΔ, Groppa Liliana; Peru: Wendy Guisela Sotelo Diaz, Jose Eduardo Chavez Corrales, Mijahil Pavel Cornejo Ortega, Manuel Gustavo Leon Portocarrero, Risto Perich CamposΔ, Luis Rasec Morales Olazabal; Philippines: Michael Tee*Δ, James Bermas/Joseph Antigua, Juan Javier Lichauco, Edgar RamiterreΔ; Poland: Marek Brzosko, Przemysław Kotyla, Hanna Mastalerz, Maria Misterska-SkóraΔ, Anna Zubrzycka-Sienkiewicz; Russia: Olga Bugrova, Ludmila Evstigneeva, Svetlana SmakotinaΔ, Bagdadi Alikhanov, Tatiana Raskina, Tatiana Kropotina, Ludmila Savina; South Korea: Chul-Soo Cho, Kichul Shin; Taiwan: James Cheng-Chung Wei, Chung-Ming Huang; Thailand: Surasak Nilganuwong; Tunisia: Bassem LouzirΔ, Fatma Boussema, Neirouz Ghannouchi, Zouhir Bahloul; USA: Emily Jane Herron Box, Aikaterini Thanou*Δ, Alvaro Ocampo, Kimberly Smith, Cummins Lue. The Steering Committee also thank the Neovacs team, who oversaw the global project management (chief medical officer: Thérèse Croughs*Δ; project managers: Stéphane Ouary*, Sandrine Kahi* and Allia Gati*; study coordinators: Marion Laborie* and Lisa Lambert; head of regulatory affairs: Nathalie Thomas-Pujol; chief scientific officer: Géraldine Grouard-Vogel; pharmaceutical development director: Olivier Dhellin), the adjudication committee members for the British Isles Lupus Assessment Group review, the independent data safety monitoring board members and all vendors who contributed to the study: WCT: global project management; MCT/Novotech: locoregional monitoring; UBC: pharmacovigilance; Q2 Solutions: central laboratory; EUROFINS: Immune response (ADME), IFN gene signature (AROS); CROS-NT: data management, eCRF/IWRS; VENN Life Sciences: statistics; Fisher: IMP (packaging and distribution); Sunnikan and Cyclop consulting: audit plan and conduct.

Contributors FH drafted the manuscript and AT, MT and TC were involved in revising it critically for important content. All other authors were involved in revising the article prior the submission. FH had full access to all data in the study and takes the responsibility for the integrity of the data and the accuracy of the data analyses. FH, AT, MT and TC were responsible for study conception and design. All authors, except TC, were involved in data acquisition, while FH, AT, MT and TC were responsible for data analysis and interpretation.

Funding This study was funded by Neovacs S.A.

Competing interests FH reports grants from Neovacs during the conduct of the study and outside the submitted work. AT reports personal fees from Neovacs during the conduct of the study. MT reports grants and personal fees from Neovacs during the conduct of the study. TC was an employee and shareholder of Neovacs at the time of the conduct and analyses of the data.

Patient consent for publication Not required.

Ethics approval Argentina: CEIC Larrea; ID: 1254/63/2015. Belgium: CEC Hospitalo-Facultaire; ID: 2015/12JUI/339. Chile: (1) CE-Cientifico Adultos, ID: EC692186/15; (2) Comite de Evaluacion Etico-Cientifico, ID: 13/August/2015. Colombia: (1) CEI Riesgo de Fractura, ID: 22 959; (2) CE en el Área de la Salud, ID: #129; (3) CEIC-CAIMED; ID: CEIC-CAIMED/#016/15; (4) Comité de Investigaciones y Ética en investigación Hospital Pablo Tobón Uribe, ID: #09/2016; (5) Comité de Ética en Investigación del Oriente, ID: #4. Croatia: CEC, ID:530-07/15-01/47. France: CPP Nord-Ouest IV, ID: CPP 15/35. Georgia: (1) EC of Ltd 'National Institute of Endocrinology', ID: Letter of 19/June/2015; (2) EC of Ltd 'Medulla', Chemotherapy and Immunotherapy, ID: Letter of 5/June/2015; (3) EC of Ltd 'Medicore', ID: Letter of 04/June/2015; (4) EC of 'Mediclub', ID: Letter of 02/June/2015. Germany: (1) EC der Justus Liebig-Universität, ID: 122/15; (2) EC des Landes Berlin, ID: 122/15; (3) Geschäftsstelle Ethikkommission Universität zu Köln Gebäude, ID:122/15; (4) EC der Landesärztekammer Rheinland-Pfalz, ID:122/15; (5) EC bei der LMU München, ID: 122/15; (6) Medizinische Hochschule Hannover Ethikkommission, ID: 1122/15. Italy: (1) Comitato Etico Milano, ID: 50 372 of 28/October/2015; (2) Nucleo Ricerca Clinica AOUI di Verona, ID: 688CESC; (3) Comitato di Bioetica/ Bioethics Committee, ID: E_20160020098; (4) Comitato Etico dell'Università "La Sapienza" di Roma, ID: 2912/15 (ref. 3890); (5) Segreteria Scientifica CE per la Sperimentazione clinica della Provincia di Padova, ID: 3662/AO/15; (6) CE-AVEC, ID:180/2015/0/Sper; (7) CE di Area Vasta Nord Oves, ID: 754/2015. Mexico: (1) CE en Investigación de la Unidad Medica de Diagnostico y Tratamiento S.A., ID: Control 006; (2) CE en Investigación Clinbor, ID: N° 06/2015; (3) CECEI BAC, ID: CECEI BAC 21/August/15; (4) CE en Investigación de la Clínica Bajío CLINBA, S, ID: 15–034; (5) MCCR, ID: MCCR-15/1213 (v) Comité de Investigación del Antiguo Hospital Civil de Guadalajara, ID:234/15; (6) CE en Investigación de la Unidad Medica de Diagnostico y Tratamiento S.A de C.V, ID: Letter of 06/Aug/2015; (7) CE en Investigación del Hospital Aranda de la Parra, ID: Letter of 20/August/2015; (8) CE en Investigación de la Clínica Bajío CLINBA, ID:15–034. Peru: (1) CIEI-USMP-CCM, ID: #1030–2015; (2) CIEI-USMP-CCM, ID: 1031–2015; (3) CIEI-USMP-CCM, ID: 1029–2015; (4) Comité de Etica en Investigacion del Hospital Nacional Guillermo Almenara Irigoyen, ID: Letter N°245 CEI-OCID-G-RAA-ESSALUD-2015; (5) Comité Institucional de Etica en Investigación de la Universidad de San Martin de Porres, ID: 1032–2015-CIEI-USMP-CCM. Philipines: (1) EC of Medical Center of Manila; ID: MMERC 2015–05; (2) EC of Chong Hua Hospital, ID: IRB-2615–08; (3) EC of St. Luke's Medical Centre, ID: CT-15074; (4) EC of Southern Philippines Medical Centre, ID: P15091101. Poland: CEC Komisja Bioetyczna przy Okręgowej Radzie Lekarskiej Wielkopolskiej Izby Lekarskiej, ID: Poland 121/2015. Russia: (1) Kemerovo Regional Clinical Hospital LEC, ID: SBEI HPE OSMU/#123; (2) CEC: Ministry of Healthcare of the Russian Federation, Ethics LEC of State Budgetary Healthcare Institution, ID: 23; (3) Ministry of Healthcare of the Russian Federation, Ethics Council-Ethics Committee of State Budgetary Healthcare Institution of Sverdlovsk Region 'Sverdlovsk Regional Clinical Hospital #1', ID: RSEC/#117; (4) CEC: Ministry of Healthcare of the Russian Federation, Ethics Council; LEC: EC of State Budgetary Educational Institution of Higher Professional Education 'Saratov State Medical University', ID: EC of SBEI Saratov SMU/# 1; (5) Local Independent Ethics and Evidence of Medical Scientific Trials Committee of State Autonomous Healthcare Institution of Kemerovo Region 'Kemerovo Regional Clinical Hospital', ID: SAHI/KRCH/ #1; (6) EC of Central Clinical Hospital of Russian Academy of Sciences CCH RAS, ID: Ethics Committee of CCH/RAS/#92; (7) EC of State Budget Healthcare Institution of Kemerovo region 'Regional Clinical Hospital for War Veterans', ID: #103/2; (8) LEC of North-Western State Medical University, ID: LEC of SBE HPENWSMU Mechnikov/#8; (9) LEC of 'Medical-Sanitary Unit #157', ID: #7; (10) Russia EC of Budgetary Healthcare Institution of Omsk Region 'Regional Clinical Hospital', ID: EC of BHIOR'RCH'/#15; (11) Russia LEC of 'Clinical-Diagnostic Center ' Ultramed, ID: #4. South Korea: (1) IRB of The Catholic University, ID: SC15BGGS0112; (2) IRB of Boramae Medical Centre, ID: 20150714/16-2015-97/081; (3) IRB of Eulji University Hospital, ID: 2015-07-012-001; (4) IRB of Konkuk University Medical Centre, ID: KUH1010693; (5) IRB of Ajou University Hospital, ID: AJIRB-MED-BDR-15–302. Spain: CE de la Investigación con Medicamentos, ID :Act N° 392. Switzerland: Commission cantonale (VD) d'éthique de la recherche sur l'être humain, ID: 388/15. Taiwan: (1) REC, China Medical University Hospital, ID: 104-REC2-127; (2) IRB, ID: 1-105-05-019; (3) IRB, Taipei Veterans General Hospital, ID: 2016–02 004BU; (4) IRB, Chang Gung Medical Foundation, ID: 104-5919A; (5) Research Ethics Committee, National Taiwan University Hospital, ID: 201 602 031MSD; (6) IRB, Chung Shan Medical University Hospital, ID: CS2-15110. Thailand: (1) Siriraj IRB, ID :450/2558; (2) Research EC3, ID: 499/2015; (3) CREC, Rajavithi Hospital, ID: 025/58BPs. Moldova: Comitetul Național de Etică pentru studiul clinic al medicamentelor și metodelor noi de tratament, ID: 199. USA: (1) Schulman/ Advarra IRB, ID: Pro00025952; (2) Columbia Research Human Research Protection Office IRB, ID: IRB-AAAQ8645; (3) University of California, San Diego USCD-HRPP, ID: 170 273. Tunisia: CPP SE prêtant à l'expérimentation médicale ou scientifique des médicaments destinés à la médecine humaine, ID: TN2016-INT-IND-31.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Additional study design details are available on clinical trial.gov (NCT02665364) and clinicaltrialsregister.eu (EudraCT 2015-001341-86). Per regulation, main study results will be posted on the same registries.

ORCID iD

Frederic A Houssiau http://orcid.org/0000-0003-1451-083X

REFERENCES

1 Bengtsson AA, Sturfelt G, Truedsson L, *et al*. Activation of type I interferon system in systemic lupus erythematosus correlates with disease activity but not with antiretroviral antibodies. *Lupus* 2000;9:664–71.
2 Blanco P*et al*. Induction of dendritic cell differentiation by IFN-alpha in systemic lupus erythematosus. *Science* 2001;294:1540–3.
3 Bennett L, Palucka AK, Arce E, *et al*. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. *J Exp Med* 2003;197:711–23.
4 McBride JM, Jiang J, Abbas AR, *et al*. Safety and pharmacodynamics of rontalizumab in patients with systemic lupus erythematosus: results of a phase I, placebo-controlled, double-blind, dose-escalation study. *Arthritis Rheum* 2012;64:3666–76.
5 Kalunian KC, Merrill JT, Maciuca R, *et al*. A phase II study of the efficacy and safety of rontalizumab (rhuMAb interferon-α) in patients with systemic lupus erythematosus (rose). *Ann Rheum Dis* 2016;75:196–202.
6 Khamashta M, Merrill JT, Werth VP, *et al*. Sifalimumab, an anti-interferon-α monoclonal antibody, in moderate to severe systemic lupus erythematosus: a randomised, double-blind, placebo-controlled study. *Ann Rheum Dis* 2016;75:1909–16.
7 Furie R, Khamashta M, Merrill JT, *et al*. Anifrolumab, an Anti-Interferon-α receptor monoclonal antibody, in moderate-to-severe systemic lupus erythematosus. *Arthritis & Rheumatology* 2017;69:376–86.
8 Furie RA, Morand EF, Bruce IN, *et al*. Type I interferon inhibitor anifrolumab in active systemic lupus erythematosus (TULIP-1): a randomised, controlled, phase 3 trial. *The Lancet Rheumatology* 2019;1:e208–19.
9 ACR/ARP. Efficacy and safety of Anifrolumab in patients with moderate to severe systemic lupus erythematosus: results of the second phase 3 randomized controlled trial: Abstract L17, 2019. Available: https://acrabstracts.org/abstract/efficacy-and-safety-of-anifrolumab-in-patients-with-moderate-to-severe-systemic-lupus-erythematosus-results-of-the-second-phase-3-randomized-controlled-trial/
10 Wallace DJ, Furie RA, Tanaka Y, *et al*. Baricitinib for systemic lupus erythematosus: a double-blind, randomised, placebo-controlled, phase 2 trial. *The Lancet* 2018;392:222–31.
11 Mathian A, Amoura Z, Adam E, *et al*. Active immunisation of human interferon α transgenic mice with a human interferon α Kinoid induces antibodies that neutralise interferon α in sera from patients with systemic lupus erythematosus. *Ann Rheum Dis* 2011;70:1138–43.
12 Zagury D, Le Buanec H, Mathian A, *et al*. IFN kinoid vaccine-induced neutralizing antibodies prevent clinical manifestations in a lupus flare murine model. *Proc Natl Acad Sci U S A* 2009;106:5294–9.
13 Lauwerys BR, Hachulla E, Spertini F, *et al*. Down-Regulation of interferon signature in systemic lupus erythematosus patients by active immunization with interferon α-kinoid. *Arthritis Rheum* 2013;65:447–56.
14 Ducreux J, Houssiau FA, Vandepapelière P, *et al*. Interferon α kinoid induces neutralizing anti-interferon α antibodies that decrease the expression of interferon-induced and B cell activation associated transcripts: analysis of extended follow-up data from the interferon α kinoid phase I/II study. *Rheumatology* 2016;55:1901–5.
15 Isenberg DA, Rahman A, Allen E, *et al*. BILAG 2004. development and initial validation of an updated version of the British Isles lupus assessment group's disease activity index for patients with systemic lupus erythematosus. *Rheumatology* 2005;44:902–6.
16 Hochberg MC. Updating the American College of rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1997;40.
17 Gladman DD, Ibañez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. *J Rheumatol* 2002;29:288–91.
18 Wallace DJ, Kalunian K, Petri MA, *et al*. Efficacy and safety of epratuzumab in patients with moderate/severe active systemic lupus erythematosus: results from EMBLEM, a phase IIb, randomised, double-blind, placebo-controlled, multicentre study. *Ann Rheum Dis* 2014;73:183–90.

19 Navarra SV, Guzmán RM, Gallacher AE, *et al*. Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomised, placebo-controlled, phase 3 trial. *The Lancet* 2011;377:721–31.
20 Gladman D, Ginzler E, Goldsmith C, *et al*. The development and initial validation of the systemic lupus international collaborating clinics/American College of rheumatology damage index for systemic lupus erythematosus. *Arthritis Rheum* 1996;39:363–9.
21 Albrecht J, Taylor L, Berlin JA, *et al*. The CLASI (cutaneous lupus erythematosus disease area and severity index): an outcome instrument for cutaneous lupus erythematosus. *J Invest Dermatol* 2005;125:889–94.
22 Franklyn K, Lau CS, Navarra SV, *et al*. Definition and initial validation of a lupus low disease activity state (LLDAS). *Ann Rheum Dis* 2016;75:1615–21.
23 Ware JE, Sherbourne CD. The mos 36-item short-form health survey (SF-36). I. conceptual framework and item selection. *Med Care* 1992;30:473–83.
24 Yao Y, Higgs BW, Morehouse C, *et al*. Development of potential pharmacodynamic and diagnostic markers for Anti-IFN-α monoclonal antibody trials in systemic lupus erythematosus. *Hum Genomics Proteomics* 2009;1.
25 Tsang-A-Sjoe MWP, Bultink IEM, Heslinga M, *et al*. Both prolonged remission and lupus low disease activity state are associated with reduced damage accrual in systemic lupus erythematosus. *Rheumatology* 2017;56:121–8.
26 Golder V, Kandane-Rathnayake R, Huq M, *et al*. Lupus low disease activity state: a prospective validation study. *Lancet Rheumatology* 2019.
27 Golder V, Kandane-Rathnayake R, Hoi AY-B, *et al*. Association of the lupus low disease activity state (LLDAS) with health-related quality of life in a multinational prospective study. *Arthritis Res Ther* 2017;19.
28 Yeo AL, Koelmeyer R, Kandane-Rathnayake R, *et al*. Lupus low disease activity state is associated with reduced direct healthcare costs in patients with systemic lupus erythematosus. *Arthritis Care Res* 2019.
29 Petri M, Purvey S, Fang H, *et al*. Predictors of organ damage in systemic lupus erythematosus: the Hopkins lupus cohort. *Arthritis Rheum* 2012;64:4021–8.
30 van Vollenhoven RF, Mosca M, Bertsias G, *et al*. Treat-to-target in systemic lupus erythematosus: recommendations from an international Task force. *Ann Rheum Dis* 2014;73:958–67.
31 Fanouriakis A, Kostopoulou M, Alunno A, *et al*. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. *Ann Rheum Dis* 2019;78:736–45.
32 Guiducci C, Gong M, Xu Z, *et al*. Tlr recognition of self nucleic acids hampers glucocorticoid activity in lupus. *Nature* 2010;465:937–41.

CLINICAL SCIENCE

Neuropsychiatric events in systemic lupus erythematosus: a longitudinal analysis of outcomes in an international inception cohort using a multistate model approach

John G Hanly ,[1] Murray B Urowitz ,[2] Caroline Gordon,[3] Sang-Cheol Bae ,[4] Juanita Romero-Diaz,[5] Jorge Sanchez-Guerrero,[2] Sasha Bernatsky ,[6] Ann E Clarke,[7] Daniel J Wallace,[8] David A Isenberg ,[9] Anisur Rahman,[9] Joan T Merrill,[10] Paul R Fortin ,[11] Dafna D Gladman ,[2] Ian N Bruce,[12] Michelle Petri ,[13] Ellen M Ginzler,[14] Mary Anne Dooley,[15] Rosalind Ramsey-Goldman,[16] Susan Manzi,[17] Andreas Jönsen,[18] Graciela S Alarcón,[19] Ronald F van Vollenhoven,[20] Cynthia Aranow ,[21] Meggan Mackay,[21] Guillermo Ruiz-Irastorza,[22] Sam Lim,[23] Murat Inanc,[24] Kenneth C Kalunian,[25] Søren Jacobsen,[26] Christine A Peschken,[27] Diane L Kamen,[28] Anca Askanase,[29] Vernon Farewell[30]

Handling editor Josef S Smolen

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-216150).

For numbered affiliations see end of article.

Correspondence to
Dr John G Hanly, Queen Elizabeth IIHlth Svc Ctr, Halifax, NS B3H 4K4, Canada; john.hanly@cdha.nshealth.ca

Received 8 August 2019
Revised 11 December 2019
Accepted 11 December 2019
Published Online First 8 January 2020

To cite: Hanly JG, Urowitz MB, Gordon C, *et al*. *Ann Rheum Dis* 2020;**79**:356–362.

ABSTRACT

Objectives Using a reversible multistate model, we prospectively examined neuropsychiatric (NP) events for attribution, outcome and association with health-related quality of life (HRQoL), in an international, inception cohort of systemic lupus erythematosus (SLE) patients.

Methods Annual assessments for 19 NP events attributed to SLE and non-SLE causes, physician determination of outcome and patient HRQoL (short-form (SF)-36 scores) were measured. Time-to-event analysis and multistate modelling examined the onset, recurrence and transition between NP states.

Results NP events occurred in 955/1827 (52.3%) patients and 592/1910 (31.0%) unique events were attributed to SLE. In the first 2 years of follow-up the relative risk (95% CI) for SLE NP events was 6.16 (4.96, 7.66) and non-SLE events was 4.66 (4.01, 5.43) compared with thereafter. Patients without SLE NP events at initial assessment had a 74% probability of being event free at 10 years. For non-SLE NP events the estimate was 48%. The majority of NP events resolved over 10 years but mortality was higher in patients with NP events attributed to SLE (16%) versus patients with no NPSLE events (6%) while the rate was comparable in patients with non-SLE NP events (7%) compared with patients with no non-SLE events (6%). Patients with NP events had lower SF-36 summary scores compared with those without NP events and resolved NP states ($p<0.001$).

Conclusions NP events occur most frequently around the diagnosis of SLE. Although the majority of events resolve they are associated with reduced HRQoL and excess mortality. Multistate modelling is well suited for the assessment of NP events in SLE.

Key messages

What is already known about this subject?

- Involvement of the nervous system in SLE is well recognised but the frequency and outcomes have not been well documented in different stages of the disease.

What does this study add?

- All neuropsychiatric (NP) events were documented over a mean follow-up of 7.6 years in a large observational inception cohort study of SLE patients.
- A multistate modelling approach was used to describe the frequency, attribution, association with health-related quality of life and clinical outcome of NP events.
- Predictive probability models were derived to estimate the likelihood of changing NP states over the first 10 years of SLE.

INTRODUCTION

Nervous system disease in SLE consists of neurologic and psychiatric events, predominantly affecting the central nervous system.[1] Neuropsychiatric (NP) events vary in frequency, complexity, time of onset, rates of resolution and recurrence. Approximately 30% of NP events are attributed to SLE,[2 3] although the rate varies between individual manifestations. Regardless of attribution, the majority of events are associated with lower self-reported health-related quality of life (HRQoL).[2]

There are few clinical trials to guide interventions in SLE patients with NP events. Many observational studies are single centre experiences, frequently cross-sectional in design and use prevalent SLE cohorts. Longitudinal studies have not captured the bidirectional movement of patients between remissions and relapses of NP events, the duration in different NP states and likelihood of moving from one state to another over time.

Key messages

How might this impact on clinical practice or future developments?

- Most NP events in SLE patients occur early in the disease course, have a negative impact on health-related quality of life and are attributed to SLE in 30% of cases.
- Multistate modelling is well suited to the study of NPSLE and could serve as an outcome measure in clinical trials (eg, comparing the rate of transition between NPSLE states for patients receiving active treatment and comparator) and long-term observational studies (eg, duration of time spent in different NPSLE states over the course of study). It also provides the basis for economic studies of healthcare costs in SLE patients with NP events, either attributed to SLE or non-SLE causes.

Multistate models[4] offer a convenient and flexible framework to characterise changes in NP disease and provide a dynamic representation of the disease in continuous time. They also estimate time spent in different states and probabilities of being in particular states following specified time periods. Such summary inferences are more informative than models that focus on single, often dichotomous, outcomes such as the time to a specific clinical event. In common with other time-to-event modelling, these inferences are based on all follow-up data, not simply on data from subsets of patients with a specific follow-up time.

The current study used data from a large, prospective, international, disease inception cohort of SLE patients, who underwent annual assessment for NP events for up to 18 years. The overall objective was to model, over time, patient status with respect to NP events, incorporating attribution and association with HRQoL. To capture dynamic change in NP events, we adopted a reversible multistate model characterised by transition rates between states.

PATIENTS AND METHODS

Research study network

The study was conducted by the Systemic Lupus International Collaborating Clinics (SLICC),[5] a network of 52 investigators at 43 academic centres in 16 countries. Recently diagnosed SLE patients were recruited from 31 SLICC sites in Europe, Asia and North America. Data were collected per protocol at enrolment and annually ensuring data quality, management and security. This research was planned without patient involvement.

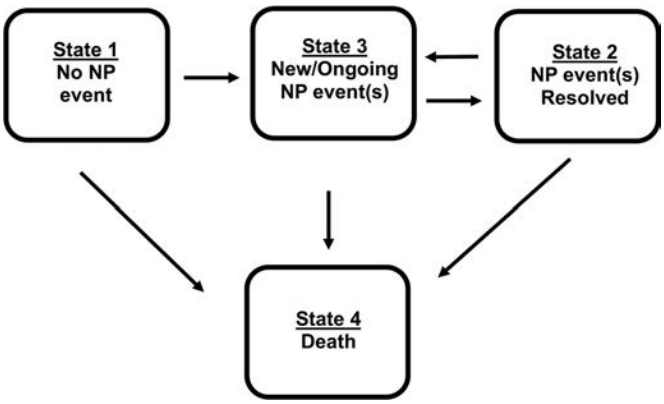


Figure 1 Reversible multistate Markovian model for observed transitions in neuropsychiatric (NP) status in patients with SLE.

Patients

Enrolment was permitted up to 15 months following diagnosis of SLE, taken as when the revised American College of Rheumatology (ACR) classification criteria[6] were first recognised. Lupus-related variables included the SLE Disease Activity Index 2000 (SLEDAI-2K)[7] and SLICC/ACR Damage Index (SDI).[8]

NP events

NP events were characterised within an enrolment window (6 months prior to the diagnosis of SLE up to the enrolment date) using ACR case definitions for 19 NP syndromes.[9] Patients were reassessed annually within a 6-month window using a detailed protocol to record information on 19 NP syndromes,[9] presence of prespecified non-SLE causes, results of appropriate investigations, medications and outcomes. New NP events that had occurred since the last study assessment and status of previous NP events were determined at each assessment. For recurring events within an assessment period, the date of the first episode was taken as the onset of the event. Additional details are provided in online supplementary file S1.

Attribution of NP events

Factors considered in the attribution decision rules included: (1) temporal onset of NP event(s) in relation to the diagnosis of SLE; (2) concurrent non-SLE factor(s), such as potential causes ('exclusions') or contributing factors ('associations') for each NP syndrome in the glossary for the ACR case definitions of NP events[9] and (3) 'common' NP events that are frequent in normal population controls as described by Ainiala *et al*.[10] These include isolated headaches, anxiety, mild depression (mood disorders failing to meet criteria for 'major depressive-like episodes'), mild cognitive impairment (deficits in less than three of the eight specified cognitive domains) and peripheral neuropathy without electrophysiological confirmation. Two attribution decision rules of different stringency (models A and B) were derived.[11 12]

Attribution model A (more stringent)

NP events attributed to SLE (1) had their onset within the enrolment window or subsequently; (2) had no 'exclusions' or 'associations' and (3) were not one of the NP events identified by Ainiala *et al*.[10]

Attribution model B (less stringent)

NP events attributed to SLE (1) had their onset within 10 years of the diagnosis of SLE and were still present within the enrolment window, or occurred subsequently; (2) had no 'exclusions' and (3) were not one of the NP events identified by Ainiala *et al*.[10]

All NP events attributed to SLE using model A were included in the NP events using model B. All other events were classified as a non-SLE NP event.[13 14]

Outcome of NP events

Physician-generated 7-point Likert scale score at each follow-up assessment compared the change in NP events between onset and follow-up (1=patient demise, 2=much worse, 3=worse, 4=no change, 5=improved, 6=much improved, 7=resolved).[15] Separately a patient-generated SF-36 questionnaire at each assessment provided eight subscale scores, and mental component summary (MCS) and physical component summary (PCS) scores.[15 16]

Table 1 Observed changes between neuropsychiatric (NP) states and death for patients with SLE NP events and non-SLE NP events as determined using attribution model B. The lower part of the table shows the estimated average (95% CI) of total time (years) spent in NP and death* states over 10 years of follow-up after cohort entry

	Transition to state			
Transition from state	**No NP**	**Resolved NP**	**New/ongoing NP**	**Death***
SLE NP events				
No NP	12 539	–	387	61
Resolved NP	–	1561	60	18
New/ongoing NP	–	270	1541	21
Non-SLE NP events				
No NP	8782	–	810	66
Resolved NP	–	2603	224	13
New/ongoing NP	–	645	3294	21
	Estimated time (years) spent in state			
SLE NP	8.81 (8.70 to 8.92)	0.39 (0.33 to 0.44)	0.56 (0.50 to 0.64)	0.24 (0.21 to 0.32)
Non-SLE NP	7.32 (7.14 to 7.48)	0.92 (0.84 to 1.01)	1.46 (1.36 to 1.58)	0.30 (0.24 to 0.37)

*For the lower portion of table 1, this refers to the average portion of the 10-year observation period that would be left after the death of a patient.

Statistical analysis

Two multistate patient level models were examined (figure 1), one for NP events attributed to SLE (model B) and the other for non-SLE events. Non-SLE events were ignored during modelling of SLE events and vice-versa. The four states were:

1. No NP event ever.
2. No current NP event but ≥1 in the past. State entry was the time of resolution of NP event(s).
3. New/ongoing NP event(s) with state entry at onset of NP event.
4. Death.

Modelling assumed transitions occurred at any time, not just at assessments. Each site investigator provided the approximate dates for onset and resolution of NP events and precise dates for death.

The time origin was 6 months before SLE diagnosis. The transition rate which characterises the probability of changing from State 1 to State 3 in the first 2 years of follow-up was allowed to differ from the rate thereafter, as many events occurred in the earlier time period. All other transition rates were assumed constant. Patients could move back and forth between States 2 (resolution of NP event) and 3 (new/ongoing NP event). The death rate was assumed to be the same from States 2 and 3 but a separate rate was allowed from State 1 (no history of NP events). The model can be extended to allow explanatory variables to influence transition rates through a regression model on the logarithm of the transition rates. Maximum likelihood estimation of the model was implemented using the R[17] package 'msm'.[18] SF-36 analyses used linear regression models with robust estimation via generalised estimating equations to adjust for correlation between multiple measurements for the same patient.

RESULTS

Patients

One-thousand eight hundred and twenty-seven patients were recruited from October 1999 through December 2011, from USA (n=540 (29.5%)), Europe (n=477 (26.1%)), Canada (n=418 (22.9%)), Mexico (n=223 (12.2%)) and Asia (n=169 (9.3%)). At enrolment, the mean (SD) age was 35.1 (13.3) years,

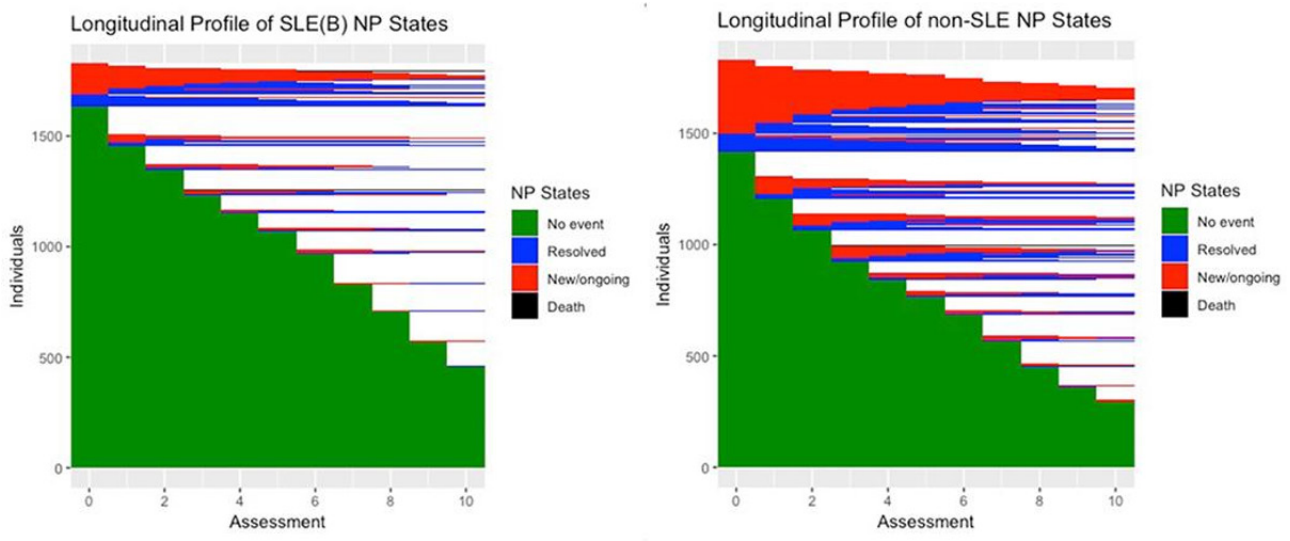


Figure 2 Layered plot of trajectories of neuropsychiatric (NP) states recorded at study assessments, for all study patients, ordered sequentially at each assessment by NP states.

88.8% of patients were female, with variable race/ethnicity (Caucasian 48.8%, African 16.8%, Hispanic 15.4%, Asian 15.1% and other 3.9%) and the mean (SD) disease duration was 5.6 (4.2) months. The mean (SD) SLEDAI-2K was 5.3 (5.4) and SDI was 0.32 (0.74). Medications at enrolment included corticosteroids (70.3%), antimalarials (67.4%), immunosuppressants (40.1%), warfarin (5.4%), low dose aspirin (14.3%), antidepressants (10.1%), anticonvulsants (4.4%) and antipsychotic drugs (0.7%). The mean follow-up was 7.6±4.6 years, with 1–19 assessments and ended in September 2017. One hundred patients died during the study.

NP manifestations

NP events occurred in 955/1827 (52.3%) patients and 493/1827 (27.0%) had ≥2 events. There were 1910 unique NP events, encompassing all 19 NP syndromes,[9] of which 1749 (91.6%) involved the central nervous system (CNS) and 161 (8.4%) the peripheral nervous system.[9] The NP events attributed to SLE varied from 17.9% (attribution model A) to 31.0% (attribution model B) and occurred in 13.5% (model A) to 21.2% (model B) of patients. Summary outcomes are provided in online supplementary tables S2 and S3 and detailed outcomes of individual manifestations may be found in previous publications.[13 14 19–22]

Transition rates between NP states, duration spent in each state and time to event analysis

The number of observed changes between NP states, for SLE and non-SLE events (table 1) provide the basis for estimation of the multistate models.

Some patients remained in the same NP state and others moved through one or more states. The table also summarises the estimated average time spent in NP states or death state over 10 years of follow-up, assuming all patients were in the no NP event state initially. A composite of NP state occupancy and duration is illustrated in lasagna plots (figure 2) where trajectories are displayed as a layered plot. For NP events attributed to SLE, the time spent in a new/ongoing NP state was lower (0.56) and the time spent in the no NP state was higher (8.81) than for NP events attributed to non-SLE causes (1.46 for new/ongoing state and 7.32 for the no NP state).

The time to onset of new and recurrent NP events is illustrated in figure 3. For recurrent events the time origin was the resolution of any previous event. The rate of occurrence of first NP events was highest in the early years following the diagnosis of SLE, consistent with our modelling strategy of allowing a differential rate of events in the first 2 years of follow-up. Based on the multistate models, the estimated relative risk (RR, 95% CI) in the first 2 years compared with the subsequent time period was 6.16 (4.96, 7.66) for SLE NP events and 4.66 (4.01, 5.43) for non-SLE events. Recurrent SLE NP events occurred at a higher rate than first events with little difference in these rates for non-SLE events.

Predictive probabilities for transitioning between NP states over time

The same multistate models provide estimates of the probability of having changed NP states or died over defined periods of time. The estimates for 10 years after entry into the no NP, resolved NP or new/ongoing NP event states are summarised in table 2.

For patients with no SLE NP event at cohort entry, there was an estimated 74% probability of remaining free of such events 10 years later. For NP events not attributed to SLE the estimate was 48%. For patients with resolved or new/ongoing NP events, the majority would be in a resolved NP state after 10 years. Estimates of having new/ongoing NP events after 10 years of follow-up were 13% and 20% for SLE NP events if the patient started follow-up in the resolved and new/ongoing states, respectively, and 26% and 31% for non-SLE NP events. Of note, the estimates of dying in a 10-year period were higher in patients with NP events (new/ongoing or resolved) attributed to SLE (16%) versus patients without SLE NP events (6%) with an estimated RR (95% CI) of death of 4.3 (2.7, 6.7). Patients with non-SLE NP events had a lower death rate (7%), similar to patients with no non-SLE events (6%), with an estimated RR of 1.3 (0.8, 2.0). After adjustment for age and postsecondary education, the RR associated with SLE NP events was slightly lower (RR=2.9 (1.6, 5.6)) but still substantially higher than that for non-SLE events (RR=1.3 (0.7, 2.4)).

A comprehensive investigation of predictors for onset and resolution of NP events at the patient level is beyond the scope of this paper. Preliminary analyses in online supplementary tables S4 and S5 provide RR estimates, based on a multistate model using multivariate regression for predictors at cohort entry. The effects are assumed to be the same on the transitions to the NP event state from both the no NP event and the resolved state. The results suggest that males have higher rates of onset and resolution for both types of NP events. Asian race and postsecondary education are protective for SLE NP event onset, and

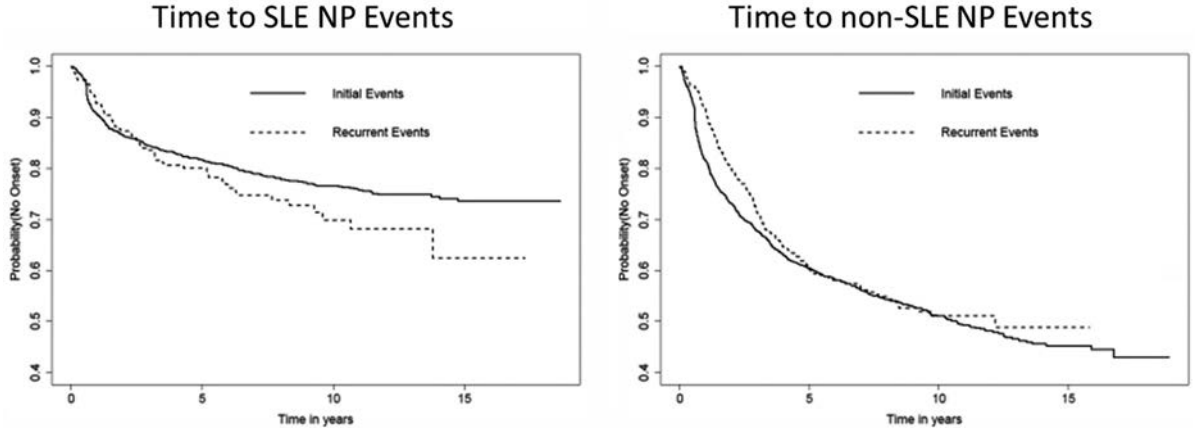


Figure 3 Time from diagnosis of SLE to onset of new and recurrent SLE neuropsychiatric (NP) events (left panel) and for non-SLE NP events identified by attribution model B (right panel). For recurrent events the time origin was the resolution of any previous event. The rate of recurrent SLE NP events occurred at a higher rate than first events with little difference in these rates for non-SLE events.

Table 2 Estimated predictive probabilities of the state a patient will be in after 10 years of follow-up for the three possible initial states, tabulated separately for SLE neuropsychiatric (NP) events and non-SLE NP events as determined using attribution model B

	Estimated probability (95% CI) of being in NP state			
Initial NP state	**No NP**	**Resolved NP**	**New/ongoing NP**	**Death**
No NP				
SLE NP events	0.74 (0.72 to 0.76)	0.13 (0.12 to 0.15)	0.07 (0.06 to 0.09)	0.06 (0.05 to 0.07)
Non-SLE NP events	0.48 (0.46 to 0.51)	0.26 (0.24 to 0.28)	0.20 (0.18 to 0.21)	0.06 (0.06 to 0.06)
Resolved NP				
SLE NP events	–	0.72 (0.68 to 0.74)	0.13 (0.10 to 0.16)	0.16 (0.12 to 0.21)
Non-SLE NP events	–	0.67 (0.64 to 0.70)	0.26 (0.23 to 0.29)	0.07 (0.05 to 0.09)
New/ongoing NP				
SLE NP events	–	0.64 0.61 to 0.68)	0.20 (0.17 to 0.24)	0.16 (0.12 to 0.21)
Non-SLE NP events	–	0.62 (0.59 to 0.65)	0.31 (0.28 to 0.34)	0.07 (0.05 to 0.09)

Asian and Hispanic races are associated with somewhat higher rates of resolution for SLE events. Hispanic race and a higher age at diagnosis are protective for the onset of non-SLE NP events while all other races have a higher resolution rate for these events than Caucasians. Higher age at diagnosis is linked to a lower resolution rate.

The association between NP states and HRQoL

Patient-generated PCS and MCS scores associated with different NP states are summarised in figure 4. There were clinically lower PCS and MCS scores in the new/ongoing NP state compared with the no NP and resolved NP states (global p value <0.001). This was true for patient states, which could vary over time, defined by both SLE and non-SLE NP events.

DISCUSSION

Heterogeneity of clinical manifestations, uncertainty in causal attribution, incomplete understanding of pathogenesis and few controlled clinical trials to guide treatment contribute to the challenge of NPSLE. Long-term studies of representative SLE cohorts provide information on the clinical course with current standard of care. A multistate modelling approach determined the frequency of NP events, their clinical outcomes and impact on HRQoL in an international prospective cohort of SLE patients. Estimates for occurrence and resolution of NP events over time in patients receiving standard of care provide a benchmark for clinical trials of new therapies.

The SLICC inception cohort has information on the occurrence and outcome of individual NP manifestations in addition to other SLE manifestations, comorbidities and treatment. Our study captures the totality of nervous system events and attribution over a mean follow-up of 7.6 years. Major organ manifestations of SLE frequently present early in the disease course, and thus a disease inception cohort is advantageous. Additional benefits include standardised, comprehensive assessments for NP events and centralised data driven decisions on attribution to SLE and non-SLE causes. The attribution of up to 30% of NP events to SLE in up to 20% of SLE patients is consistent

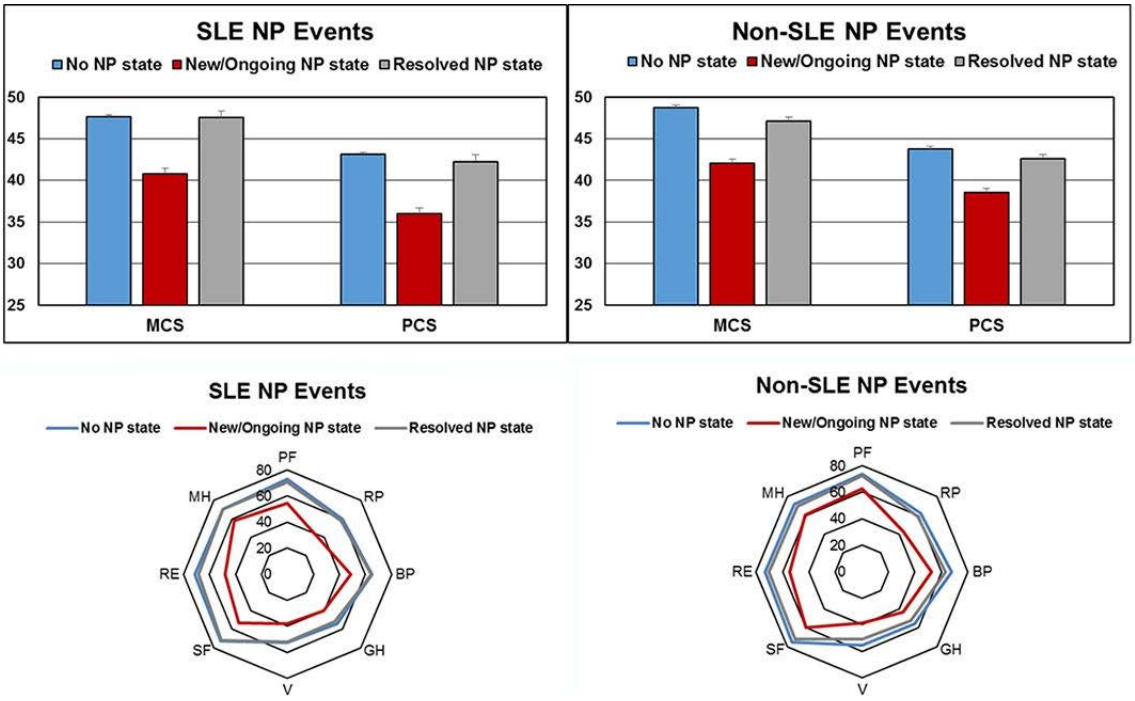


Figure 4 SF-36 physical component summary (PCS) and mental component summary (MCS) scores (mean and 95% CI) for patients in different neuropsychiatric (NP) states (global p value <0.001). The SLE NP events and non-SLE NP events were determined using attribution model B. GH, general health; BP, bodily pain; MH, mental health; PF, physical function; RE, role emotion; RP, role physical; SF, social function; V, vitality.

with previous SLICC cohort studies[12 23] and with other recent studies.[3 24]

Multistate modelling has been applied to psoriatic arthritis,[25 26] lupus nephritis[27] and organ damage accrual in SLE[28] but has not previously been used to study NPSLE. Although NP events can present or recur at any time in the disease course, they are most frequent in the first few years following SLE diagnosis. This is the case for both NP events attributed to SLE and non-SLE causes, although the rate of recurrence compared with the rate for the initial event is higher for NP events attributed to SLE. As is the case with other SLE manifestations,[27 29] the occurrence of NP manifestations has implications for the subsequent disease course. Patients whose initial state was free of SLE NP events, had a 74% likelihood of being free of SLE NP events at 10 years of follow-up. For patients still free of SLE NP events 2 years after the diagnosis, this probability rises to 84%. Thus, if patients remain event free during this time, there is a high likelihood that NPSLE manifestations will not occur subsequently. For patients whose initial SLE NP state was 'resolved', or who subsequently had all NP events resolve, there was a 72% probability of being free of SLE NP events after 10 years of follow-up.

Individual types of NP events, regardless of attribution, are associated with reduced HRQoL[12 15 23] that is clinically and statistically significant. The current study found a similar association between NP events in total and reduced HRQoL that reverts towards normal with resolution of the events. More ominously, following 10 years of follow-up, there is a higher probability (16%) of death for patients who experience SLE NP events compared with those without SLE NP events (6%) with little increase associated with non-SLE NP events (7%). Another recent large, long-term study[3] found that patients with NPSLE had a threefold higher mortality with a hazard ratio (95% CI) of 3.09 (0.03–9.21). Thus, major organ involvement by SLE carries a higher mortality risk over time, although the cause of death is not necessarily attributed to affected organ systems.[29]

Transition rates between NP states were derived from clinically meaningful changes identified by treating physicians and supported by patient self-report health status. Potential applications of multistate modelling include using transition rates as a primary outcome in clinical trials and projecting the cost of care for NP disease. For example, the estimated probability of an SLE NP event resolving within 2 years is 0.31. A clinical trial to detect a 50% improvement in that rate would require a sample size (with an alpha level of 0.05% and 80% power) of 282 patients (141/group). Furthermore, by determining the actual costs for each NP state and knowing the projected proportion of patients and the duration of time in each state, one can predict the costs of care.

There are limitations to the current study. First although an inception cohort study is well suited to document NP events occurring early and likely due to active lupus, our study is not well positioned to detect NP events later in the disease course such as stroke and cognitive impairment from atherosclerosis and vascular dementia. Further follow-up will be required to address this. Second, SLICC comprised predominantly of academic centres with a special interest in SLE that may not reflect community clinical practice. Third, as this is an observational study conducted in multiple international centres with annual study assessments there is potential for variability in data collection. Close communication between SLICC sites and operational rules for data collection and attribution have been implemented to keep this to a minimum. Finally, the complexity of a reversible multistate model to predict long-term probabilities of state occupancy requires the use of parametric assumptions for transition rates, generally that the rate of transitioning out of a state is constant and does not depend on how long a patient has been in a state.

Despite these limitations our study provides a comprehensive overview of NP events, their attribution and outcome in the first decade following the diagnosis of SLE in a representative group of SLE patients. Future studies will examine in detail the predictors of transition between the NP states and the economic costs associated with NPSLE.

Author affiliations
[1]Division of Rheumatology, Department of Medicine and Department of Pathology, Queen Elizabeth II Health Sciences Centre and Dalhousie University, Halifax, Nova Scotia, Canada
[2]Center for Prognosis Studies in the Rheumatic Diseases, Toronto Western Hospital, University of Toronto, Toronto, Ontario, Canada
[3]Rheumatology Research Group, Institute of Inflammation and Ageing, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK
[4]Department of Rheumatology, Hanyang University Hospital for Rheumatic Diseases, Seoul, Republic of Korea
[5]Instituto Nacional de Ciencias Medicas y Nutricion, Mexico City, Mexico
[6]Divisions of Rheumatology and Clinical Epidemiology, Department of medicine, McGill University, Montreal, Quebec, Canada
[7]Divisions of Rheumatology, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada
[8]Cedars-Sinai/David Geffen School of Medicine at UCLA, Los Angeles, CA, USA
[9]Centre for Rheumatology Research, Department of Medicine, University College, London, UK
[10]Department of Clinical Pharmacology, Oklahoma Medical Research Foundation, Oklahoma City, OK, USA
[11]Division of Rheumatology, Department of Medicine, CHU de Québec, Université Laval, Quebec City, Quebec, Canada
[12]Arthritis Research UK Epidemiology Unit, Faculty of Biology Medicine and Health, Manchester Academic Health Sciences Centre, The University of Manchester, and NIHR Manchester Musculoskeletal Biomedical Research Centre, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK
[13]Department of Rheumatology, Johns Hopkins University School of Medicine, Baltimore, MD, USA
[14]Department of Medicine, SUNY Downstate Medical Center, Brooklyn, NY, USA
[15]Thurston Arthritis Research Centre, University of North Carolina, Chapel Hill, NC, USA
[16]Northwestern University and Feinberg School of Medicine, Chicago, IL, USA
[17]Lupus Center of Excellence, Allegheny Health Network, Pittsburgh, PA, USA
[18]Department of Clinical Sciences Lund, Rheumatology, Lund University, Lund, Sweden
[19]Department of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA
[20]Department of Rheumatology and Clinical Immunology, University Medical Centres, Amsterdam, The Netherlands
[21]Feinstein Institute for Medical Research, Manhasset, NY, USA
[22]Autoimmune Diseases Research Unit, Department of Internal Medicine, BioCruces Bizkaia Health Research Institute, Hospital Universitario Cruces, University of the Basque Country, Barakaldo, Spain
[23]Emory University, Department of Medicine, Division of Rheumatology, Atlanta, Georgia, USA
[24]Division of Rheumatology, Department of Internal Medicine, Istanbul Medical Faculty, Istanbul University, Istanbul, Turkey
[25]UCSD School of Medicine, La Jolla, CA, USA
[26]Copenhagen Lupus and Vasculitis Clinic, 4242, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark
[27]University of Manitoba, Winnipeg, Manitoba, Canada
[28]Medical University of South Carolina, Charleston, South Carolina, USA
[29]Hospital for Joint Diseases, NYU, Seligman Centre for Advanced Therapeutics, New York, NY, USA
[30]MRC Biostatistics Unit, Cambridge Institute of Public Health, University of Cambridge, Cambridge, UK

Acknowledgements The corresponding author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive license (or non-exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd and its Licensees to permit this article (if accepted) to be published in Annals of the Rheumatic Diseases and any other BMJPGL products to exploit all subsidiary rights, as set out in our license (http://group.bmj.com/products/journals/instructions-for-authors/licence-forms).

Contributors All authors contributed to drafting and reviewing the manuscript and approved the final version for publication. JGH has access to all data and analyses. Study conception and design: JGH, MU, CG, S-CB, JR-D, JS-G, SB, AEC, DJW, DI, AR, JTM, PRF, DDG, INB, MP, EMG, MAD, RR-G, SM, GSA, RvV, CA, KK, VF. Acquisition of data, and analysis and interpretation of data: JGH, MU, CG, S-CB, JR-D, JS-G, SB, AEC, DJW, DI, AR, JTM, PRF, DDG, INB, MP, EMG, MAD, RR-G, SM, AJ, GSA, RvV, CA, MM, GR-I, SL, MI, KK, SJ, CP, DLK, AA, VF.

Funding Core funding for this investigator-initiated study was provided to JGH by the Canadian Institutes of Health Research (grant MOP-88526). Other sources of funding supported activities at individual SLICC sites: S-CB's work was supported in part by NRF-2017M3A9B4050335, Republic of Korea. CG is supported by Lupus UK, Sandwell and West Birmingham Hospitals NHS Trust and the National Institute for Health Research (NIHR)/Wellcome Trust Birmingham Clinical Research Facility. The views expressed are those of the authors(s) and not necessarily those of the NHS, the NIHR or the Department of Health. The Hopkins Lupus Cohort is supported by the NIH (grant AR43727 and 69572). The Montreal General Hospital Lupus Clinic is partially supported by the Singer Family Fund for Lupus Research. AEC holds The Arthritis Society Chair in Rheumatic Diseases at the University of Calgary. PRF holds a tier 1 Canada Research Chair on Systemic Autoimmune Rheumatic Diseases at Université Laval. INB is a National Institute for Health Research (NIHR) Senior Investigator and is supported by Arthritis Research UK, the NIHR Manchester Biomedical Centre and the NIHR/Wellcome Trust Manchester Clinical Research Facility. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health. SJ is supported by the Danish Rheumatism Association (A3865) and the Novo Nordisk Foundation (A05990). RR-G's work was supported by the NIH (grants 5UL1TR001422-02, formerly 8UL1TR000150 and UL-1RR-025741, K24-AR-02318, and P60AR064464 formerly P60-AR-48098). MAD's work was supported by the NIH grant RR00046. GR-I is supported by the Department of Education, Universities and Research of the Basque Government. DI and AR are supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre. SL's work was supported, in part, by the Centers for Disease Control and Prevention grant U01DP005119.

Competing interests RvV has received grants from BMS, GSK, Lilly, Pfizer, UCB Pharma, personal fees from AbbVie, AstraZeneca, Biotest, Celgene, GSK, Janssen, Lilly, Novartis, Pfizer, Servier, UCB, outside the submitted work.

Patient consent for publication Not required.

Ethics approval The Nova Scotia Health Authority central zone Research Ethics Board and each of the participating centres' research ethics review boards approved the study.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information. Not applicable as all data relevant to the study are included in the article of in supplementary files.

ORCID iDs
John G Hanly http://orcid.org/0000-0003-1029-9483
Murray B Urowitz http://orcid.org/0000-0001-7506-9166
Sang-Cheol Bae http://orcid.org/0000-0003-4658-1093
Sasha Bernatsky http://orcid.org/0000-0002-9515-2802
David A Isenberg http://orcid.org/0000-0001-9514-2455
Paul R Fortin http://orcid.org/0000-0002-7278-2596
Dafna D Gladman http://orcid.org/0000-0002-9074-0592
Michelle Petri http://orcid.org/0000-0003-1441-5373
Cynthia Aranow http://orcid.org/0000-0001-9299-0053

REFERENCES

1 Hanly JG, Kozora E, Beyea SD, *et al*. Review: nervous system disease in systemic lupus erythematosus: current status and future directions. *Arthritis Rheumatol* 2019;71:33–42.
2 Hanly JG. Diagnosis and management of neuropsychiatric SLE. *Nat Rev Rheumatol* 2014;10:338–47.
3 Ahn GY, Kim D, Won S, *et al*. Prevalence, risk factors, and impact on mortality of neuropsychiatric lupus: a prospective, single-center study. *Lupus* 2018;27:1338–47.
4 Cook RJ, Lawless JF. *Multistate models for the analysis of life history data*. Boca Raton, Florida, USA: CRC Press, 2018.
5 Isenberg D, Ramsey-Goldman R, SLICC Group. Systemic Lupus International Collaborating Group--onwards and upwards? *Lupus* 2006;15:606–7.
6 Hochberg MC. Updating the American College of rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1997;40:1725.
7 Gladman DD, Ibañez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. *J Rheumatol* 2002;29:288–91.
8 Gladman D, Ginzler E, Goldsmith C, *et al*. The development and initial validation of the systemic lupus international collaborating Clinics/American College of rheumatology damage index for systemic lupus erythematosus. *Arthritis Rheum* 1996;39:363–9.
9 ACR AD HOC COMMITTEE ON NEUROPSYCHIATRIC LUPUS NOMENCLATURE. The American College of rheumatology Nomenclature and case definitions for neuropsychiatric lupus syndromes. *Arthritis Rheum* 1999;42:599–608.
10 Ainiala H, Hietaharju A, Loukkola J, *et al*. Validity of the new American College of rheumatology criteria for neuropsychiatric lupus syndromes: a population-based evaluation. *Arthritis Rheum* 2001;45:419–23.
11 Hanly JG, Urowitz MB, Su L, *et al*. Short-term outcome of neuropsychiatric events in systemic lupus erythematosus upon enrollment into an international inception cohort study. *Arthritis Rheum* 2008;59:721–9.
12 Hanly JG, Urowitz MB, Sanchez-Guerrero J, *et al*. Neuropsychiatric events at the time of diagnosis of systemic lupus erythematosus: an international inception cohort study. *Arthritis Rheum* 2007;56:265–73.
13 Hanly JG, Li Q, Su L, *et al*. Psychosis in systemic lupus erythematosus: results from an international inception cohort study. *Arthritis Rheumatol* 2019;71:281–9.
14 Hanly JG, Li Q, Su L, *et al*. Peripheral nervous system disease in systemic lupus erythematosus: results from an international inception cohort study. *Arthritis Rheumatol* 2019;29.
15 Hanly JG, Urowitz MB, Jackson D, *et al*. SF-36 summary and subscale scores are reliable outcomes of neuropsychiatric events in systemic lupus erythematosus. *Ann Rheum Dis* 2011;70:961–7.
16 Thumboo J, Fong KY, Ng TP, *et al*. Validation of the MOS SF-36 for quality of life assessment of patients with systemic lupus erythematosus in Singapore. *J Rheumatol* 1999;26:97–102.
17 R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2015.
18 Jackson CH. Multi-State Models for Panel Data: The msm Package for *R*. *J Stat Softw* 2011;38:1–29.
19 Hanly JG, Li Q, Su L, *et al*. Cerebrovascular events in systemic lupus erythematosus: results from an international inception cohort study. *Arthritis Care Res (Hoboken)* 2018;70:1478–87.
20 Hanly JG, Su L, Urowitz MB, *et al*. Mood disorders in systemic lupus erythematosus: results from an international inception cohort study. *Arthritis Rheumatol* 2015;67:1837–47.
21 Hanly JG, Urowitz MB, O'Keeffe AG, *et al*. Headache in systemic lupus erythematosus: results from a prospective, International inception cohort study. *Arthritis Rheum* 2013;65:2887–97.
22 Hanly JG, Urowitz MB, Su L, *et al*. Seizure disorders in systemic lupus erythematosus results from an international, prospective, inception cohort study. *Ann Rheum Dis* 2012;71:1502–9.
23 Hanly JG, Li Q, Su L, *et al*. Cerebrovascular events in systemic lupus erythematosus. *Arthritis Care Res* 2018.
24 Fanouriakis A, Pamfil C, Rednic S, *et al*. Is it primary neuropsychiatric systemic lupus erythematosus? performance of existing Attribution models using physician judgment as the gold standard. *Clin Exp Rheumatol* 2016;34:910–7.
25 Farewell VT, Su L. A multistate model for events defined by prolonged observation. *Biostatistics* 2011;12:102–11.
26 Jackson CH, Su L, Gladman DD, *et al*. On modelling minimal disease activity. *Arthritis Care Res (Hoboken)* 2016;68:388–93.
27 Hanly JG, Su L, Urowitz MB, *et al*. A longitudinal analysis of outcomes of lupus nephritis in an international inception cohort using a multistate model approach. *Arthritis Rheumatol* 2016;68:1932–44.
28 Bruce IN, O'Keeffe AG, Farewell V, *et al*. Factors associated with damage accrual in patients with systemic lupus erythematosus: results from the systemic lupus international collaborating clinics (SLICC) inception cohort. *Ann Rheum Dis* 2015;74:1706–13.
29 Hanly JG, O'Keeffe AG, Su L, *et al*. The frequency and outcome of lupus nephritis: results from an international inception cohort study. *Rheumatology (Oxford)* 2016;55:252–62.

CLINICAL SCIENCE

High genetic risk score is associated with early disease onset, damage accrual and decreased survival in systemic lupus erythematosus

Sarah Reid ,[1] Andrei Alexsson,[1] Martina Frodlund,[2] David Morris,[3] Johanna K Sandling,[1] Karin Bolin,[1] Elisabet Svenungsson ,[4] Andreas Jönsen,[5] Christine Bengtsson,[6] Iva Gunnarsson,[4] Vera Illescas Rodriguez,[4] Anders Bengtsson,[5] Sabine Arve ,[5] Solbritt Rantapää-Dahlqvist ,[6] Maija-Leena Eloranta,[1] Ann-Christine Syvänen,[7] Christopher Sjöwall ,[2] Timothy James Vyse,[3] Lars Rönnblom ,[1] Dag Leonard [1]

Handling editor Josef S Smolen

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-216227).

For numbered affiliations see end of article.

Correspondence to
Professor Lars Rönnblom, Rheumatology and Science for Life Laboratory, Department of Medical Sciences, Uppsala University, 752 36 Uppsala, Sweden;
lars.ronnblom@medsci.uu.se

Results included have been presented at the 13th International Congress on SLE (Reid et al., Lupus Sci Med 2019;6:A154-5.) and at the ACR/ARHP Meeting (Reid et al., Arthritis Rheum 2017; 69 (suppl 10).

Received 27 August 2019
Revised 18 November 2019
Accepted 25 November 2019
Published Online First 11 December 2019

To cite: Reid S, Alexsson A, Frodlund M, *et al*. *Ann Rheum Dis* 2020;**79**:363–369.

ABSTRACT

Objectives To investigate associations between a high genetic disease risk and disease severity in patients with systemic lupus erythematosus (SLE).

Methods Patients with SLE (n=1001, discovery cohort and n=5524, replication cohort) and healthy controls (n=2802 and n=9859) were genotyped using a 200K Immunochip single nucleotide polymorphism array. A genetic risk score (GRS) was assigned to each individual based on 57 SLE risk loci.

Results SLE was more prevalent in the high, compared with the low, GRS-quartile (OR 12.32 (9.53 to 15.71), $p=7.9\times10^{-86}$ and OR 7.48 (6.73 to 8.32), $p=2.2\times10^{-304}$ for the discovery and the replication cohorts, respectively). In the discovery cohort, patients in the high GRS-quartile had a 6-year earlier mean disease onset (HR 1.47 (1.22 to 1.75), $p=4.3\times10^{-5}$), displayed higher prevalence of damage accrual (OR 1.47 (1.06 to 2.04), $p=2.0\times10^{-2}$), renal disorder (OR 2.22 (1.50 to 3.27), $p=5.9\times10^{-5}$), anti-dsDNA (OR 1.83 (1.19 to 2.81), $p=6.1\times10^{-3}$), end-stage renal disease (ESRD) (OR 5.58 (1.50 to 20.79), $p=1.0\times10^{-2}$), proliferative nephritis (OR 2.42 (1.30 to 4.49), $p=5.1\times10^{-3}$), anti-cardiolipin-IgG (OR 1.89 (1.13 to 3.18), $p=1.6\times10^{-2}$), anti-β_2-glycoprotein-I-IgG (OR 2.29 (1.29 to 4.06), $p=4.8\times10^{-3}$) and positive lupus anticoagulant test (OR 2.12 (1.16 to 3.89), $p=1.5\times10^{-2}$) compared with patients in the low GRS-quartile. Survival analysis showed earlier onset of the first organ damage (HR 1.51 (1.04 to 2.25), $p=3.7\times10^{-2}$), first cardiovascular event (HR 1.65 (1.03 to 2.64), $p=2.6\times10^{-2}$), nephritis (HR 2.53 (1.72 to 3.71), $p=9.6\times10^{-7}$), ESRD (HR 6.78 (1.78 to 26.86), $p=6.5\times10^{-3}$) and decreased overall survival (HR 1.83 (1.02 to 3.30), $p=4.3\times10^{-2}$) in high to low quartile comparison.

Conclusions A high GRS is associated with increased risk of organ damage, renal dysfunction and all-cause mortality. Our results indicate that genetic profiling may be useful for predicting outcomes in patients with SLE.

Key messages

What is already known about this subject?

- The field of genetics has been revolutionised by genome-wide association studies, with over 100 genetic loci associated with systemic lupus erythematosus (SLE) discovered.
- Genetic risk scores have shown promise for understanding the polygenic contribution to many complex diseases but have been scarcely investigated in SLE.

What does this study add?

- In the present study, we demonstrate that a high genetic risk is associated with an early onset of SLE, increased organ damage, cardiovascular disease and end-stage renal disease, as well as impaired survival.

How might this impact on clinical practice or future developments?

- Our results suggest that genetic profiling of patients with SLE may be useful for predicting outcome of the disease.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a chronic disease characterised by loss of tolerance to self-antigens, formation of immune complexes and an activated type I interferon system.[1–3] Despite improved prognosis, the mortality rate stills exceeds that of the general population.[4] Due to active inflammation, prolonged corticosteroid use, comorbidities and factors unrelated to SLE, organ damage accumulates in the majority of patients over time,[1 5 6] with cardiovascular disease and renal failure being strong risk factors for premature mortality.[4 7–9]

Familial aggregation and twin studies provide compelling evidence of genetic predisposition in SLE, with a more than 10-fold higher concordance rate for monozygotic than for dizygotic twins.[10 11] The genetic aetiology is complex, with single nucleotide polymorphisms (SNPs) at more than 100 genetic loci associated with SLE identified at genome-wide significance.[1 12–16] While susceptibility to SLE appears to increase with the number of these risk loci,[13] specific disease manifestations

may be associated with a subset of polymorphisms. For example, variants of Signal Transducer and Activator of Transcription 4 (*STAT4*), have displayed association with nephritis, ischaemic stroke, severe renal insufficiency and a younger age at disease onset[17–20] as well as an increased overall risk of organ damage.[21] For the majority of SLE susceptibility loci however, no links to specific disease subphenotypes have been demonstrated.

Comprehension of the genetic contribution to permanent organ damage is important for understanding the pathogenesis of SLE. Additionally, prediction of disease outcome is essential for optimising monitoring and treatment strategies, to reduce both unnecessary side-effects and long-term disease complications. Genetic risk scores (GRSs) have been applied in several fields of medicine, and studies have demonstrated their ability to predict matters like cardiovascular disease, prostate cancer risk and body mass index scores.[22–24] In SLE, few studies have assessed the relationship between the cumulative genetic risk and disease subphenotypes,[25–28] and the association between the polygenic risk and disease severity is unknown. In this study, we examined the relationship between a high GRS and clinical manifestations associated with more severe SLE phenotypes, including organ damage, defined by the Systemic Lupus Collaborating Clinics (SLICC)/American College of Rheumatology (ACR) Damage Index (SDI),[29] cardiovascular events (CVE) and end-stage renal disease (ESRD).

PATIENTS, HEALTHY INDIVIDUALS AND METHODS

Patients and healthy controls

The discovery cohort included 1001 patients from the University clinics in Uppsala, Linköping, Karolinska Institute (Stockholm), Lund, and from the four northern-most counties in Sweden. All subjects fulfilled ≥4 ACR-82 classification criteria for SLE and were of European descent.[30] Clinical data were collected from the patients' medical files, including SDI scores,[29] the ACR-82 classification criteria, clinical antiphospholipid syndrome (APS) diagnosis, glomerular filtration rate, chronic kidney disease (CKD) stages, ESRD, renal biopsy data and CVE, defined as myocardial infarction, ischaemic cerebrovascular disease or venous thromboembolism (VTE). For definitions, see online supplementary file 1. Patient characteristics are summarised in table 1. For prevalences of SDI scores per organ domain, see online supplementary table 1. Control individuals were healthy blood donors from Uppsala (Uppsala Bioresource) and Lund or population based controls from Stockholm and the four northernmost counties of Sweden. The replication cohort included 5524 patients with SLE and 9859 healthy controls of European ancestry, defined by principal component analysis, described in Langefeld *et al*.[1]

Genotyping and construction of the genetic risk score

Genotyping of the discovery cohort was performed using the Illumina 200K Immunochip SNP array by the SNP&SEQ Technology platform at Science for Life Laboratory in Uppsala, Sweden. For quality control (QC) procedures, see online supplementary file.

Cumulative GRSs were assigned to each individual based on SNPs with previous association with SLE at genome wide significance in the European population from the publication by Chen *et al*.[13] The inclusion criteria (see online supplementary file 1) allowed for inclusion of 57 SNPs (online supplementary table 2). For each SNP, the natural logarithm of the OR for SLE susceptibility based on comparisons between the 1001 patients and 2802 controls in the discovery cohort was multiplied by the number of risk alleles in each individual. The sum of all products for each patient was defined as the GRS. In addition, a risk allele count (RAC) of the 57 SNPs in each individual was performed by adding the total number of risk alleles. Finally, an HLA-GRS was constructed, see online supplementary file 1 and online supplementary table 3.

Individuals in the replication cohort were independently genotyped using the Illumina 200K Immunochip SNP array, available at https://www.ebi.ac.uk/gwas/. A RAC and GRS was assigned to each patient and control using the same 57 SNPs and OR as in the discovery cohort analysis, see online supplementary table 2. Individuals included in the discovery cohort analysis or with <100% genotype success rate of the 57 SNPs were excluded from the replication cohort (pi HAT >0.9). For genotyping and QC procedures of the replication cohort, see Langefeld *et al*.[1]

Statistical analysis

We used ordinal or logistic regression to assess differences in prevalences between groups. Age was included as a covariate in all analyses, and significant results were subsequently analysed in a second model, with the age at SLE diagnosis as an additional covariate. The generalised Wilcoxon test was employed to assess differences in survival. For more information on statistical analysis, see online supplementary file. Statistical analyses were performed using R.[31] Unadjusted $p<0.05$ were considered statistically significant.

RESULTS

Genetic characteristics of patients and healthy individuals

Initially, we performed a RAC in each individual in the discovery cohort and as can be seen in figure 1A, the RAC followed a Gaussian distribution, with higher mean scores in patients than in healthy controls (mean (SD) 52.71 (4.81) compared with 48.95 (4.71)). The prevalence of SLE was higher in individuals with a RAC in the highest, compared with the lowest, quartile (OR 7.81 (6.19–9.85), $p=1.9\times10^{-67}$). To test whether the difference between groups would increase when considering the contribution to SLE by each SNP, a weighted GRS was constructed. Similar to the RAC, the GRS followed a Gaussian distribution with higher mean scores in patients than in controls (mean (SD) 8.52 (1.20) compared with 7.45 (1.20)) (figure 1B). In the discovery cohort, the probability that an individual had SLE increased with increasing GRS (figure 1C) and was significantly higher in the highest, compared with the lowest, GRS-quartile (OR 12.32 (9.53 to 15.71), $p=7.9\times10^{-86}$). Moreover, patients with a GRS in the high quartile received their SLE diagnosis significantly earlier in life, with a mean age at SLE onset in the high and low quartiles of 33 and 39 years, respectively (figure 1D).

We subsequently employed receiver operating characteristic (ROC) curve analysis to compare prediction accuracies of the scores. The GRS was significantly better than the RAC at discriminating between patients and controls (area under the ROC curve (AUC) 0.78 compared with 0.71, $p_{comparison}=1.4\times10^{-14}$). In addition, the prediction accuracy of the GRS was higher in patients<20 years at SLE onset ($p=3.0\times10^{-3}$ compared with patients aged 20–40 years at onset, $p=2.35\times10^{-6}$ compared with patients aged >40 years at onset) (figure 2).

Replication cohort validation

The RAC and the GRS were validated using genetic data from a replication cohort including more than 15 000 patients and controls. Results show a higher probability of SLE in the high,

Table 1 Prevalence of clinical manifestations and serology vs associations with the genetic risk score in the Discovery cohort

		GRS, high vs low quartiles		GRS, continuous	
	n (%)	OR (95% CI)*	P value†	OR (95% CI)‡	P value†
Deceased at follow-up	99 (10)	1.79 (0.93 to 3.46)	8.0×10^{-2}	**1.30 (1.07 to 1.59)**	**9.4×10^{-3}**
Male gender	132 (13)	1.27 (0.77 to 2.12)	3.4×10^{-1}	1.07 (0.91 to 1.24)	4.2×10^{-1}
SDI scores[29]		**1.47 (1.06 to 2.04)**	**2.0×10^{-2}**	**1.13 (1.03 to 1.24)**	**1.4×10^{-2}**
SLE criteria, ACR-82[30]					
Malar rash	565 (56)	0.88 (0.61 to 1.26)	5.4×10^{-1}	0.94 (0.85 to 1.05)	2.6×10^{-1}
Discoid rash	236 (24)	0.85 (0.56 to 1.30)	4.7×10^{-1}	0.94 (0.83 to 1.07)	3.4×10^{-1}
Photosensitivity	680 (68)	0.75 (0.51 to 1.09)	1.2×10^{-1}	**0.88 (0.79 to 0.99)**	**2.6×10^{-2}**
Oral ulcers	249 (25)	1.07 (0.71 to 1.62)	8.5×10^{-1}	1.02 (0.91 to 1.15)	7.0×10^{-1}
Arthritis	800 (80)	0.74 (0.47 to 1.17)	2.0×10^{-1}	0.91 (0.80 to 1.04)	1.5×10^{-1}
Serositis	447 (45)	0.95 (0.66 to 1.36)	8.2×10^{-1}	0.95 (0.86 to 1.06)	3.6×10^{-1}
Renal disorder	342 (34)	**2.22 (1.50 to 3.27)**	**5.9×10^{-5}**	**1.29 (1.16 to 1.44)**	**7.0×10^{-6}**
Neurological disorder	105 (10)	1.12 (0.77 to 1.62)	5.6×10^{-1}	1.09 (0.92 to 1.29)	3.3×10^{-1}
Haematological disorder	616 (62)	1.04 (0.87 to 1.25)	6.5×10^{-1}	1.05 (0.94 to 1.17)	3.7×10^{-1}
Immunological disorder	686 (69)	**2.03 (1.38 to 2.98)**	**3.6×10^{-4}**	**1.29 (1.15 to 1.45)**	**1.6×10^{-5}**
dsDNA antibodies	477 (62)	**1.83 (1.19 to 2.81)**	**6.1×10^{-3}**	**1.31 (1.15 to 1.50)**	**4.2×10^{-5}**
Sm antibodies	95 (13)	1.24 (0.65 to 2.37)	5.2×10^{-1}	1.10 (0.90 to 1.33)	3.5×10^{-1}
ANA	970 (98)	2.29 (0.59 to 8.89)	2.3×10^{-1}	1.37 (0.91 to 2.07)	1.4×10^{-1}
Renal biopsy data[47]					
WHO Class I-II	32 (14)	1.67 (0.61 to 4.60)	3.2×10^{-1}	1.17 (0.86 to 1.59)	3.3×10^{-1}
WHO Class III-IV	133 (60)	**2.42 (1.30 to 4.49)**	**5.1×10^{-3}**	**1.36 (1.14 to 1.62)**	**7.5×10^{-4}**
WHO Class V	31 (14)	1.88 (0.70 to 5.10)	2.1×10^{-1}	1.10 (0.80 to 1.51)	5.6×10^{-1}
Other§	20 (9)	0.95 (0.29 to 3.13)	9.5×10^{-1}	1.01 (0.68 to 1.50)	9.5×10^{-1}
CKD stages[48]		**2.16 (1.31 to 3.56)**	**2.6×10^{-3}**	**1.26 (1.09 to 1.47)**	**2.4×10^{-3}**
ESRD	24 (2)	**5.58 (1.50 to 20.79)**	**1.0×10^{-2}**	**1.65 (1.18 to 2.32)**	**3.6×10^{-3}**
Antiphospholipid antibodies					
Any aPL	257 (38)	**1.84 (1.16 to 2.9)**	**9.4×10^{-3}**	**1.15 (1.00 to 1.32)**	**4.9×10^{-2}**
Triple positive aPLs¶	119 (20)	**2.27 (1.02 to 5.09)**	**4.6×10^{-2}**	**1.30 (1.02 to 1.66)**	**3.2×10^{-2}**
LA	121 (22)	**2.12 (1.16 to 3.89)**	**1.5×10^{-2}**	**1.21 (1.02 to 1.45)**	**3.3×10^{-2}**
aCL-IgG	181 (27)	**1.89 (1.13 to 3.18)**	**1.6×10^{-2}**	1.14 (0.98 to 1.32)	9.1×10^{-2}
aCL-IgM	69 (13)	1.07 (0.5 to 2.29)	8.6×10^{-1}	1.13 (0.91 to 1.41)	2.7×10^{-1}
aβ_2GP-I-IgG	118 (18)	**2.29 (1.29 to 4.06)**	**4.8×10^{-3}**	**1.32 (1.11 to 1.58)**	**2.1×10^{-3}**
aβ_2GP-I-IgM	19 (11)	1.01 (0.98 to 1.05)	4.9×10^{-1}	0.91 (0.61 to 1.35)	6.3×10^{-1}
Clinical APS	132 (19)	1.35 (0.78 to 2.33)	2.8×10^{-1}	1.13 (0.96 to 1.34)	1.4×10^{-1}

Values in bold indicate p<0.05.
*OR for the high compared to the low GRS-quartile.
†Unadjusted.
‡OR for every increase of one point in the GRS (eg, from 6.5 to 7.5).
§Patients with biopsies displaying signs of nephritis but not meeting the criteria for any of the above classes[32] were classified as other.
¶Triple positivity for aPLs was defined as having positive tests for aCL (IgG or IgM) and aß2GP-I (IgG or IgM) and LA.
aCL, anticardiolipin; ACR, American College of Rheumatology; aβ_2GP-I, anti-β_2 Glycoprotein-I;aPL, anti-phospholipid antibody; APS, antiphospholipid syndrome; CKD, chronic kidney disease; ESRD, end-stage renal disease; GRS, genetic risk score; LA, lupus anticoagulant; SDI, SLICC Damage Index; SLE, systemic lupus erythematosus; SLICC, Systemic Lupus Collaborating Clinics.

compared with the low, quartile both for the RAC (OR 5.84 (5.23 to 6.53), $p=5.47\times10^{-213}$) and the GRS (OR 7.48 (6.73 to 8.32), $p=2.2\times10^{-304}$). Online supplementary figure 1A illustrates the correlation between GRS and prevalence of SLE in this cohort. In the replication cohort, ROC curve analysis showed AUCs of 0.68 and 0.71 for the RAC and the GRS, respectively ($p_{comparison}=2.2\times10^{-16}$) (online supplementary figure 1B).

Genetic risk score associations

Because the GRS was superior to the RAC in discriminating between patients and controls, subsequent analyses focused on this score. The GRS was analysed as a continuous variable in all regression analyses, with table 1 presenting ORs for a one-unit increase in the GRS. To simplify the interpretation of ORs, we also compared patients with a GRS in the extreme quartiles. There was no significant difference in SLE disease duration between the high and low GRS-quartiles (OR 1.00 (0.99 to 1.02), $p=6.7\times10^{-1}$).

The prevalence of organ damage, as defined by the SDI, increased with increasing GRS ($p=1.4\times10^{-2}$). Figure 3A illustrates the probability of having each individual SDI score for patients in the high, compared with the low, GRS-quartile, with 52%, 67% and 83% higher odds of having 2, 3 or ≥4 points on the index, respectively. In the survival analyses, the high and low GRS-quartiles were compared. The mean survival until the first organ damage was decreased in the high quartile ($p=3.7\times10^{-2}$), with affected individuals acquiring their first damage at a mean age of 43 years, compared with 51 years in the low GRS-quartile (table 2).

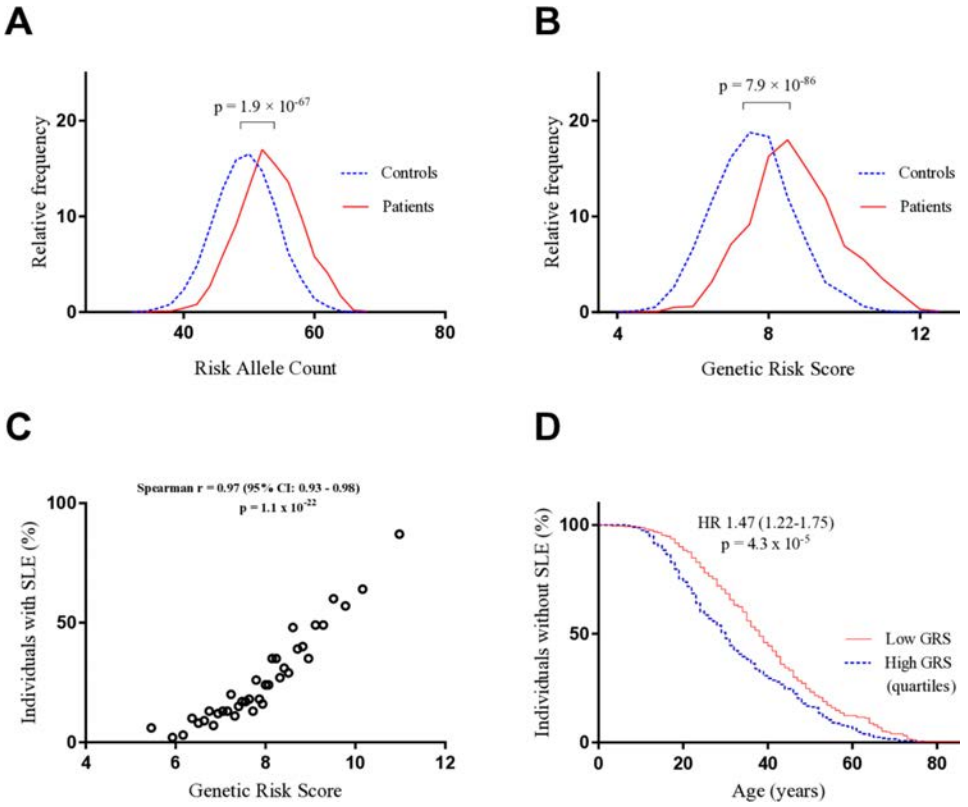


Figure 1 Cumulative genetic risk and SLE development. (A) The distribution of the RAC in the patients (n=1001) and healthy controls (n=2802). (B) The distribution of the weighted GRS in the same individuals. (C) The patients and healthy controls were ordered according to their GRSs and divided into 38 groups, each including 100 individuals (with exception of the first group, which consisted of 103 individuals). The SLE prevalence of each group was plotted against its mean GRS. (D) The survival until SLE onset was analysed for patients with a GRS in the extreme quartiles (n=500). GRS, genetic risk score; RAC, risk allele count; SLE, systemic lupus erythematosus.

Overall mortality increased with increasing GRS ($p=9.4\times10^{-3}$) (table 1) with the highest ORs for mortality observed in the groups of patients with the highest GRS (figure 3B). Patients in the high GRS-quartile further displayed a shorter mean survival compared with the low quartile ($p=4.3\times10^{-2}$) (table 2).

Because CVD is an important component of the SDI, we analysed survival until the first CVE separately. Patients in the high quartile displayed a decreased survival ($p=2.6\times10^{-2}$), with a mean age at the first event in affected individuals of 45 years, compared with 51 years in the low GRS-quartile (table 2). We subsequently divided CVE into arterial events (AE) and VTE

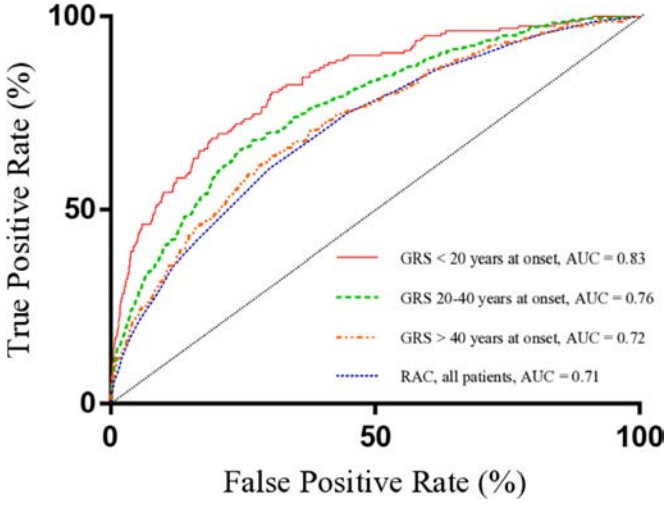


Figure 2 Prediction accuracy of the weighted GRS depending on age at SLE onset. ROC curve analysis was used to assess the prediction ability of the GRS in patients aged below 20 (n=158), 20–40 (n=475) and >40 (n=368) years at SLE diagnosis. The prediction accuracy of the unweighted RAC is shown in the same figure. AUC, area under the ROC curve; GRS, genetic risk score; RAC, risk allele count; ROC, receiver operating characteristic; SLE, systemic lupus erythematosus.

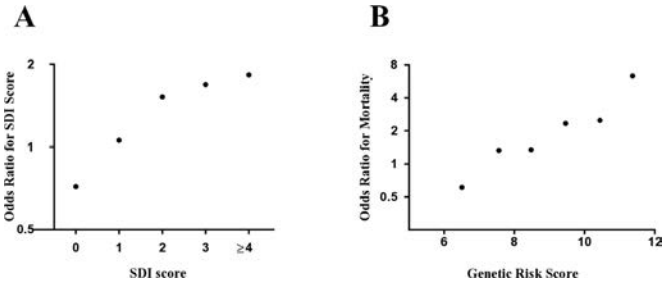


Figure 3 Association of high GRS with organ damage and overall mortality. (A) In five separate logistic regression models, the probability of having 0 vs >0, or 1/2/3/≥4 vs 0, points on the SLICC SDI was calculated for patients with a GRS in the high, compared with the low, quartile. Age was included as a covariate in the analyses. (B) Using the same statistical model and covariate as in A, the OR for mortality compared with patients with a GRS<7 was plotted for patients with a GRS of 7–8, 8–9, 9–10, 10–11 and >11. Patients with a GRS<7 were compared with patients with a GRS>7. GRS, genetic risk score, SDI, SLICC Damage Index; SLICC, Systemic Lupus Collaborating Clinics.

Table 2 Survival comparisons based on patients with a GRS in the extreme quartiles in the Discovery cohort

	N patients		Mean age at event* (mean survival†)			
	Affected	Unaffected	High quartile	Low quartile	HR (95% CI)	P value‡
First SDI score	124	92	43 (51)	51 (59)	**1.51 (1.04 to 2.25)**	**3.7×10^{-2}**
First CVE	114	308	45 (64)	51 (70)	**1.65 (1.03 to 2.64)**	**2.6×10^{-2}**
First AE	72	310	52 (69)	58 (78)	**2.16 (1.21 to 3.87)**	**9.7×10^{-3}**
First VTE	60	322	39 (75)	46 (79)	1.30 (0.78 to 2.17)	3.0×10^{-1}
Onset of ESRD	14	245	43 (82)	64 (92)	**6.78 (1.78 to 26.86)**	**6.5×10^{-3}**
Overall mortality	50	379	66 (76)	66 (82)	**1.83 (1.02 to 3.30)**	**4.3×10^{-2}**

Patients in the extreme quartiles were included as affected individuals if they met the criteria for the examined manifestation; otherwise as censored individuals.
Values in bold indicate p<0.05.
*The mean age at the event includes only affected individuals.
†The mean survival is defined as the age at which 50% of individuals in each quartile are affected by the examined event.
‡Unadjusted.
AE, arterial event (myocardial infarction or ischaemic cerebrovascular disease); CVE, cardiovascular event (AE or VTE); ESRD, end-stage renal disease; GRS, genetic risk score; SDI, SLICC Damage Index[29] ; VTE, venous thromboembolic event (deep vein thrombosis or pulmonary embolism).

and found that patients in the high GRS-quartile displayed a decreased survival until their first AE ($p=9.7\times10^{-3}$), but not their first VTE ($p=3.0\times10^{-1}$) (table 2).

Analysis of the ACR–82 criteria[30] showed that the prevalence of the renal and immunological criteria increased with increasing GRS ($p=5.9\times10^{-5}$ and $p=3.6\times10^{-4}$, respectively), with doubled odds of each manifestation in the high-to-low GRS-quartile comparison (table 1). In addition, dsDNA prevalence increased with increasing GRS (table 1). Patients in the high quartile further displayed a decreased mean survival until nephritis debut ($p=9.6\times10^{-7}$), with a mean age at nephritis onset of 31 years, compared with 39 years in the low GRS-quartile (figure 4). Next, we investigated the connection between cumulative genetics and renal dysfunction further. An increasing GRS was associated with higher stages of CKD and with development of ESRD, with five times elevated odds of ESRD in the high-to-low GRS-quartile comparison ($p=1.0\times10^{-2}$) (table 1). In addition, the mean survival until ESRD onset was decreased, with the mean onset in affected individuals occurring at 43 years in the high GRS-quartile, compared with 64 years in the low quartile (table 2). We subsequently analysed patients with positive renal biopsy results (n=222) and found that the prevalence of proliferative nephritis increased with increasing GRS (table 1).

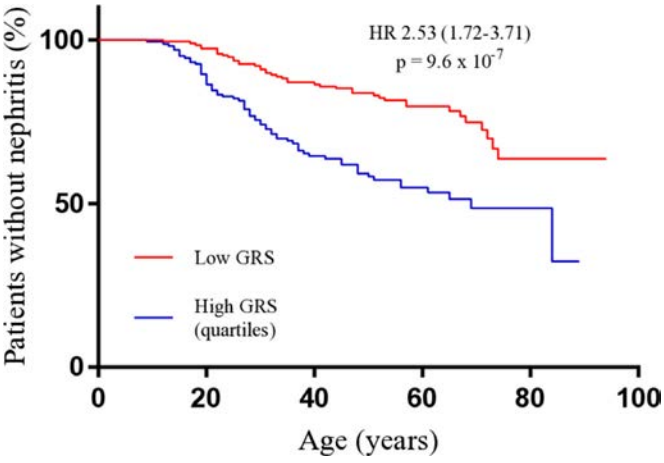


Figure 4 Survival comparison until nephritis onset in patients with a high or low GRS. Patients with a GRS in the extreme quartiles meeting the ACR-82 nephritis criterion, with a known date of nephritis diagnosis (n=109), were included as cases in the analysis, with their age at the time of nephritis diagnosis as the time variable. Patients in the extreme quartiles not meeting the nephritis criterion (n=245) were included as censored individuals, with their age at last-follow up as the time variable. The high and low quartiles were compared using the generalised Wilcoxon test. GRS, genetic risk score.

Due to the relationship between a high GRS and an earlier onset of CVE, we investigated associations between the score and the prevalence of APS/anti-phospholipid antibodies (aPLs). The GRS was not significantly associated with APS; however, patients in the high GRS-quartile were more likely to have a positive aPL test ($p=9.4\times10^{-3}$), with more than doubled odds of being triple positive (table 1). Individually, lupus anticoagulant (LA), aβ_2GP-I-IgG and aCL-IgG were significantly more prevalent in the high compared with the low quartile, with ORs of 2.12, 2.29 and 1.89, respectively (table 1).

To determine whether the association between a high GRS and early disease onset influenced other results, all previously significant associations were reanalysed with the age at SLE diagnosis included as an additional covariate. With the exception of the association between the GRS and proliferative nephritis on biopsy, all previously observed associations remained significant (online supplementary table 4).

Next, we calculated positive and negative predictive values (PPV and NPV) for our most important findings (online supplementary table 5). The GRS showed the highest predictive ability for ESRD, which at a GRS cut-off level of 9.5 had a specificity of 83%. At a prevalence of 11%,[32] the PPV and NPV were 31% and 95%, respectively.

Risk allele count, HLA-GRS and individual risk allele associations

To test whether the associations would remain when removing the weights of the GRS, all regression analyses were repeated using the unweighted RAC. With the exception of ESRD and the aPL variables, all associations remained significant (online supplementary table 6). We subsequently employed ROC curve analysis to compare prediction accuracies of the scores and found that the RAC generated a significantly better prediction of the immunological criterion[30] whereas the GRS displayed a better prediction accuracy for ESRD, aβ_2GP-I-IgG as well as presence of ≥3 aPLs (online supplementary table 6).

Next, we investigated associations between the HLA-GRS and clinical manifestations. With exception of negative associations with APS, aCL-IgM, aβ_2GP-I-IgG and LA, no significant associations were found (online supplementary table 7).

Finally, all SNPs included in the GRS were analysed individually for association with the SDI. The *STAT4* (rs11889341) and *PRDM–ATG5* (rs6568431) risk variants were associated with increased SDI scores (OR 1.29 (1.10 to 1.52), $p=2.9\times10^{-3}$ and OR 1.31 (1.11 to 1.55), $p=1.4\times10^{-3}$, respectively) whereas *TMEM39A* (rs1132200) displayed an association with lower SDI scores (OR 0.70 (0.55 to 0.90), $p=1.4\times10^{-3}$).

DISCUSSION

Our study is the first to demonstrate an association between high cumulative genetic risk and survival, organ damage, cardiovascular disease, proliferative nephritis, ESRD and antiphospholipid antibodies in patients with SLE, introducing GRSs as a potential tool for prediction of disease severity. We employed both a weighted GRS and an unweighted RAC for our analyses, and their similar prediction accuracies regarding most outcomes—including organ damage and mortality—suggest that the added effect of multiple loci plays a more central role in the contribution to disease severity than the individual contribution by any high risk SNP.

The present study confers three important findings that may aid in explaining the association of the cumulative genetic risk with organ damage. First, we demonstrate that a high GRS is associated to an earlier onset of CVE, which is an important component of the SDI.[29] Second, we found an association between a high GRS and presence of aPLs, including more than doubled odds of having a positive LA test. In addition to patients with aPLs having an increased risk of CVE,[33] the LA test has been demonstrated to be the most predictive serological test for organ damage.[34] Finally, the GRS was associated with renal involvement, higher stages of CKD, more severe biopsy classes including proliferative nephritis and, in particular, with ESRD. The renal domain is included as a separate item in the SDI, with ESRD generating more points than any other component of the index.[29] Although these variables are likely contributors to our main result, there may be other important factors associated to both the GRS and to organ damage which were not examined in this study.

Our demonstration of a 6-year difference in SLE onset between the high and low GRS-quartiles supports previous findings by both Taylor *et al*[35] and Langefelt *et al*.[1] A younger age at onset is associated with higher disease activity,[36] an increased prevalence of nephritis and prolonged corticosteroid treatment,[37] and the risk of acquiring organ damage in this group of patients is thus increased.[5 38] We therefore included the age at SLE diagnosis as an additional covariate in our regression analysis and found only a small reduction in the effect size. Thus, the association between cumulative genetics and early disease onset may only to a limited extent explain our findings.

We found two individual variants positively associated with increased organ damage. The *STAT4* variant has previously been associated with a more severe disease phenotype including ischaemic stroke and increased SDI scores.[17–21] Patients with SLE carrying this risk variant display an augmented IFN-γ production in T cells and elevated STAT1 expression in B cells.[39 40] Because of the entailed potential therapeutic opportunity, we believe our confirmation of the association of this variant with organ damage is valuable. The *ATG5* gene encodes a protein involved in autophagy.[41] Some studies have indicated that an altered function of this process increases the risk of lupus nephritis,[42] which is in turn associated with damage accrual.

In analysis of the HLA-GRS, we found a negative association with aPLs and clinical APS. The reason for this may be that the DRB1*03:01 tag SNP rs1269852, due to its high prevalence and OR for SLE in our cohort, made a substantial contribution to the total score. Patients carrying this SLE-HLA allele are less likely to carry the DRB1*04 and *13 alleles, which are associated with secondary APS.[43]

The strength of our study is the large population including more than 1000 well-characterised patients with SLE, the comprehensive collection of clinical data and the long mean disease duration, allowing for long time follow-up of damage accrual. The validation of the GRS in a population including more than 15 000 patients and controls also confirms the significance of the cumulative genetic score. There are, however, some limitations. The retrospective approach of our study may confer a falsely low difference in overall survival between patients with high and low GRS, as only patients deceased after year 2000 are included in our study population. In addition, we lacked data regarding cumulative prednisolone dose and cumulative disease activity, which are important risk factors for the development of organ damage.[5 44 45]

Despite displaying moderate accuracy in the prediction of the examined manifestations, the combination of their relatively high prevalence, their severity and the benefit of early detection indicates a clinical relevance to the GRS. For example, an ESRD screening test with a GRS cut-off level of 9.5 would generate 22% positive samples, of which 31% would develop the complication compared with 5% of negative cases. Importantly however, the present study explores a GRS weighted by ORs for SLE rather than for renal manifestations. As there are several SNPs associated specifically with lupus nephritis,[46] the method could be employed to design a nephritis-specific GRS with, plausibly, higher predictive accuracy.

In conclusion, a high GRS is associated with a more severe SLE phenotype involving an earlier onset of the disease, more organ damage and renal dysfunction, as well as impaired survival. Our results indicate that genetic profiling may provide a tool for predicting disease outcome and thus aid in the clinical decision process.

Author affiliations
[1]Rheumatology and Science for Life Laboratory, Department of Medical Sciences, Uppsala University, Uppsala, Sweden
[2]Rheumatology/Division of Neuro and Inflammation Sciences, Department of Clinical and Experimental Medicine, Linköping University, Linkoping, Sweden
[3]Department of Medical and Molecular Genetics, King's College London, London, UK
[4]Rheumatology Unit, Department of Medicine, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden
[5]Rheumatology, Department of Clinical Sciences, Lund University, Lund, Sweden
[6]Department of Public Health and Clinical Medicine/Rheumatology, Umeå University, Umeå, Sweden
[7]Molecular Medicine and Science for Life Laboratory, Department of Medical Sciences, Uppsala University, Uppsala, Sweden

Acknowledgements We thank the patients and healthy volunteers for participating, clinicians and researches contributing to the replication cohort and Cane Yaka, Rezvan Kiani, Karolina Tandre and Pascal Pucholt. This work was supported by the National Institute for Health Research Biomedical Research Centre (NIHR BRC) at Guy's and St Thomas' NHS Foundation and King's College London

Contributors SR, DL and LR designed the study. SR, MF, JKS, KB, ES, AJ, CB, IG, VIR, AB, SA, SR-D, M-LE, A-CS, CS, LR and DL collected the data. DLM and TJV collected and compiled the replication cohort data. SR, AA and DL performed the statistical analysis. SR, DL and LR analysed the data. SR, DL and LR wrote the manuscript. All authors revised the manuscript critically for important intellectual content and approved the final version of the manuscript.

Funding This project was funded by the Swedish Research Council for Medicine and Health, Knut and Alice Wallenberg Foundation, Swedish Society of Medicine and Ingegerd Johansson donation, Swedish Rheumatism Foundation, King Gustaf V's 80-year Foundation, Uppsala County Council and Uppsala University Hospital (ALF), Stockholm County Council (ALF), Uppsala University, Swedish Heart-Lung foundation, Selander Foundation, Agnes and Mac Rudberg Foundation and Gustaf

Prim Foundation. Several funders, recognised in Langefeld et al including the SLEGEN consortium, supported and funded the replication cohort data.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The study protocol was approved by the Regional Ethical Review Board, Uppsala (DNR 2016/155) and the local ethics committees. All participants gave their informed consent.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

ORCID iDs
Sarah Reid http://orcid.org/0000-0003-4065-6875
Elisabet Svenungsson http://orcid.org/0000-0003-3396-3244
Sabine Arve http://orcid.org/0000-0002-3347-5550
Solbritt Rantapää-Dahlqvist http://orcid.org/0000-0001-8259-3863
Christopher Sjöwall http://orcid.org/0000-0003-0900-2048
Lars Rönnblom http://orcid.org/0000-0001-9403-6503
Dag Leonard https://orcid.org/0000-0002-6275-7282

REFERENCES

1 Langefeld CD, Ainsworth HC, Cunninghame Graham DS, *et al*. Transancestral mapping and genetic load in systemic lupus erythematosus. *Nat Commun* 2017;8:16021.
2 Lo MS, Tsokos GC. Recent developments in systemic lupus erythematosus pathogenesis and applications for therapy. *Curr Opin Rheumatol* 2018;30:222–8.
3 Rönnblom L, Leonard D. Interferon pathway in SLE: one key to unlocking the mystery of the disease. *Lupus Sci Med* 2019;6:e000270.
4 Bernatsky S, Boivin J-F, Joseph L, *et al*. Mortality in systemic lupus erythematosus. *Arthritis Rheum* 2006;54:2550–7.
5 Bruce IN, O'Keeffe AG, Farewell V, *et al*. Factors associated with damage accrual in patients with systemic lupus erythematosus: results from the systemic lupus international collaborating clinics (SLICC) inception cohort. *Ann Rheum Dis* 2015;74:1706–13.
6 Frodlund M, Reid S, Wetterö J, *et al*. The majority of Swedish systemic lupus erythematosus patients are still affected by irreversible organ impairment: factors related to damage accrual in two regional cohorts. *Lupus* 2019;28:1261–72.
7 Björnådal L, Yin L, Granath F, *et al*. Cardiovascular disease a hazard despite improved prognosis in patients with systemic lupus erythematosus: results from a Swedish population based study 1964-95. *J Rheumatol* 2004;31:713–9.
8 Almaani S, Meara A, Rovin BH. Update on lupus nephritis. *CJASN* 2017;12:825–35.
9 Gustafsson JT, Simard JF, Gunnarsson I, *et al*. Risk factors for cardiovascular mortality in patients with systemic lupus erythematosus, a prospective cohort study. *Arthritis Res Ther* 2012;14.
10 Deapen D, Escalante A, Weinrib L, *et al*. A revised estimate of twin concordance in systemic lupus erythematosus. *Arthritis Rheum* 1992;35:311–8.
11 Sestak AL, Fürnrohr BG, Harley JB, *et al*. The genetics of systemic lupus erythematosus and implications for targeted therapy. *Ann Rheum Dis* 2011;70:i37–43.
12 Bentham J, Morris DL, Graham DSC, *et al*. Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. *Nat Genet* 2015;47:1457–64.
13 Chen L, Morris DL, Vyse TJ. Genetic advances in systemic lupus erythematosus. *Curr Opin Rheumatol* 2017;29:423–33.
14 Harley JB, Alarcón-Riquelme ME, Criswell LA, *et al*. Genome-Wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. *Nat Genet* 2008;40:204–10.
15 Hom G, Graham RR, Modrek B, *et al*. Association of Systemic Lupus Erythematosus with *C8orf13–BLK* and *ITGAM–ITGAX*. *N Engl J Med* 2008;358:900–9.
16 Morris DL, Sheng Y, Zhang Y, *et al*. Genome-Wide association meta-analysis in Chinese and European individuals identifies ten new loci associated with systemic lupus erythematosus. *Nat Genet* 2016;48:940–6.
17 Ceccarelli F, Perricone C, Borgiani P, *et al*. Genetic factors in systemic lupus erythematosus: contribution to disease phenotype. *J Immunol Res* 2015;2015:1–11.
18 Sigurdsson S, Nordmark G, Garnier S, *et al*. A risk haplotype of *STAT4* for systemic lupus erythematosus is over-expressed, correlates with anti-dsDNA and shows additive effects with two risk alleles of IRF5. *Hum Mol Genet* 2008;17:2868–76.
19 Svenungsson E, Gustafsson J, Leonard D, *et al*. A STAT4 risk allele is associated with ischaemic cerebrovascular events and anti-phospholipid antibodies in systemic lupus erythematosus. *Ann Rheum Dis* 2010;69:834–40.
20 Bolin K, Sandling JK, Zickert A, *et al*. Association of STAT4 polymorphism with severe renal insufficiency in lupus nephritis. *PLoS One* 2013;8:e84450.
21 Nageeb RS, Omran AA, Nageeb GS, *et al*. Stat4 gene polymorphism in two major autoimmune diseases (multiple sclerosis and juvenile onset systemic lupus erythematosus) and its relation to disease severity. *Egypt J Neurol Psychiatr Neurosurg* 2018;54.
22 Belsky DW, Moffitt TE, Sugden K, *et al*. Development and evaluation of a genetic risk score for obesity. *Biodemography Soc Biol* 2013;59:85–100.
23 Krarup NT, Borglykke A, Allin KH, *et al*. A genetic risk score of 45 coronary artery disease risk variants associates with increased risk of myocardial infarction in 6041 Danish individuals. *Atherosclerosis* 2015;240:305–10.
24 Helfand B. A comparison of genetic risk score with family history for estimating prostate cancer risk. *Asian J Androl* 2016;18:515–9.
25 Hughes T, Adler A, Merrill JT, *et al*. Analysis of autosomal genes reveals gene–sex interactions and higher total genetic risk in men with systemic lupus erythematosus. *Ann Rheum Dis* 2012;71:694–9.
26 Gianfrancesco MA, Balzer L, Taylor KE, *et al*. Genetic risk and longitudinal disease activity in systemic lupus erythematosus using targeted maximum likelihood estimation. *Genes Immun* 2016;17:358–62.
27 Webber D, Cao J, Dominguez D, *et al*. Association of systemic lupus erythematosus (SLE) genetic susceptibility loci with lupus nephritis in childhood-onset and adult-onset SLE. *Rheumatology* 2019;17.
28 Alonso-Perez E, Suarez-Gestal M, Calaza M, *et al*. Lack of replication of higher genetic risk load in men than in women with systemic lupus erythematosus. *Arthritis Res Ther* 2014;16.
29 Gladman D, Ginzler E, Goldsmith C, *et al*. The development and initial validation of the systemic lupus international collaborating Clinics/American College of rheumatology damage index for systemic lupus erythematosus. *Arthritis Rheum* 1996;39:363–9.
30 Tan EM, Cohen AS, Fries JF, *et al*. The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1982;25:1271–7.
31 Team RC. R: a language and environment for statistical computing: R foundation for statistical computing, 2018. Available: https://www.R-project.org/
32 Tektonidou MG, Dasgupta A, Ward MM. Risk of end-stage renal disease in patients with lupus nephritis, 1971-2015: a systematic review and Bayesian meta-analysis. *Arthritis Rheumatol* 2016;68:1432–41.
33 Habe K, Wada H, Matsumoto T, *et al*. Presence of antiphospholipid antibodies as a risk factor for thrombotic events in patients with connective tissue diseases and idiopathic thrombocytopenic purpura. *Intern Med* 2016;55:589–95.
34 Petri M, Purvey S, Fang H, *et al*. Predictors of organ damage in systemic lupus erythematosus: the Hopkins lupus cohort. *Arthritis Rheum* 2012;64:4021–8.
35 Taylor KE, Chung SA, Graham RR, *et al*. Risk alleles for systemic lupus erythematosus in a large case-control collection and associations with clinical subphenotypes. *PLoS Genet* 2011;7:e1001311.
36 Sousa S, Gonçalves MJ, Inês LS, *et al*. Clinical features and long-term outcomes of systemic lupus erythematosus: comparative data of childhood, adult and late-onset disease in a national register. *Rheumatol Int* 2016;36:955–60.
37 Tomic-Lucic A, Petrovic R, Radak-Perovic M, *et al*. Late-Onset systemic lupus erythematosus: clinical features, course, and prognosis. *Clin Rheumatol* 2013;32:1053–8.
38 Alarcón GS, Roseman JM, McGwin G, *et al*. Systemic lupus erythematosus in three ethnic groups. XX. damage as a predictor of further damage. *Rheumatology* 2004;43:202–5.
39 Hagberg N, Joelsson M, Leonard D, *et al*. The *STAT4* SLE risk allele rs7574865[T] is associated with increased IL-12-induced IFN-γ production in T cells from patients with SLE. *Ann Rheum Dis* 2018;77:1070–7.
40 Patel ZH, Lu X, Miller D, *et al*. A plausibly causal functional lupus-associated risk variant in the STAT1-STAT4 locus. *Hum Mol Genet* 2018;27:2392–404.
41 Liu X, Qin H, Xu J. The role of autophagy in the pathogenesis of systemic lupus erythematosus. *Int Immunopharmacol* 2016;40:351–61.
42 Wang L, Law H. The role of autophagy in lupus nephritis. *Int J Mol Sci* 2015;16:25154–67.
43 Lundström E, Gustafsson JT, Jönsen A, *et al*. HLA-DRB1*04/*13 alleles are associated with vascular disease and antiphospholipid antibodies in systemic lupus erythematosus. *Ann Rheum Dis* 2013;72:1018–25.
44 Becker-Merok A, Nossent HC. Damage accumulation in systemic lupus erythematosus and its relation to disease activity and mortality. *J Rheumatol* 2006;33:1570–7.
45 Zonana-Nacach A, Barr SG, Magder LS, *et al*. Damage in systemic lupus erythematosus and its association with corticosteroids. *Arthritis Rheum* 2000;43:1801–8.
46 Iwamoto T, Niewold TB. Genetics of human lupus nephritis. *Clin Immunol* 2017;185:32–9.
47 Churg J, Bernstein J, Glassrock R. *Renal disease: classification and atlas of glomerular diseases*. 2nd edn. New York: Igaku-Shoin, 1995: 151–5.
48 National Kidney Foundation. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Am J Kidney Dis* 2002;39:S1–266.

CLINICAL SCIENCE

Haemodynamic phenotypes and survival in patients with systemic sclerosis: the impact of the new definition of pulmonary arterial hypertension

Panagiota Xanthouli,[1,2] Suzana Jordan,[3] Nicklas Milde,[1,2] Alberto Marra,[4] Norbert Blank,[5] Benjamin Egenlauf,[1,2] Matthias Gorenflo,[6] Satenik Harutyunova,[1,2] Hanns-Martin Lorenz,[5] Christian Nagel,[1,2,7] Vivienne Theobald,[1,2] Mona Lichtblau,[8] Charlotte Berlier,[8] Silvia Ulrich,[8] Ekkehard Grünig,[1,2] Nicola Benjamin ,[1,2] Oliver Distler [3]

Handling editor Josef S Smolen

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-216476).

For numbered affiliations see end of article.

Correspondence to
Nicola Benjamin, Centre for Pulmonary Hypertension, Thoraxklinik-Heidelberg gGmbH, Heidelberg 69126, Germany; nicola.benjamin@med.uni-heidelberg.de

PX and SJ are joint first authors. NB and OD are joint senior authors.

Received 15 October 2019
Revised 25 November 2019
Accepted 25 November 2019
Published Online First 9 December 2019

To cite: Xanthouli P, Jordan S, Milde N, *et al*. *Ann Rheum Dis* 2020;**79**:370–378.

ABSTRACT

Background In this study, we investigated the impact of the new haemodynamic definition of pulmonary arterial hypertension (PAH) as proposed by the 6th PH World Symposium on phenotypes and survival in patients with systemic sclerosis (SSc).

Methods In SSc patients who were prospectively and consecutively screened for PAH including right heart catheterisation in Heidelberg or Zurich, haemodynamic and clinical variables have been reassessed according to the new PAH definition. Patients have been followed for 3.7±3.7 (median 3.4) years; Kaplan-Meier survival analysis was performed. Patients with significant lung or left heart disease were excluded from comparative analyses.

Results The final dataset included 284 SSc patients, 146 patients (49.2%) had mean pulmonary arterial pressure (mPAP) ≤20 mm Hg, 19.3% had mPAP 21–24 mm Hg and 29.4% had mPAP ≥25 mm Hg. In the group of mildly elevated mPAP, only four patients (1.4% of the whole SSc cohort) had pulmonary vascular resistance (PVR) values ≥3 Wood Units (WU) and could be reclassified as manifest SSc-APAH. Twenty-eight (9.8%) patients with mPAP of 21–24 mm Hg and PVR ≥2 WU already presented with early pulmonary vascular disease with decreased 6 min walking distance (6MWD) ($p<0.001$), TAPSE ($p=0.004$) and pulmonary arterial compliance ($p<0.001$). A PVR ≥2 WU was associated with reduced long-term survival ($p=0.002$). PVR and 6MWD were independent prognostic predictors in multivariate analysis.

Conclusion The data of this study show that a PVR threshold ≥3 WU is too high to enable an early diagnosis of PAH. A PVR threshold ≥2 WU was already associated with pulmonary vascular disease, significantly reduced survival and would be more appropriate in SSc patients with mild PAH.

Key messages

What is already known about this subject?

► Patients with systemic sclerosis-associated pulmonary arterial hypertension present with severely impaired survival rates. The new haemodynamic classification of pulmonary hypertension will lead to changes in classification of manifest disease and possibly treatment decisions.

What does this study add?

► This study shows that patients with mildly elevated mean pulmonary arterial pressure and pulmonary vascular resistance (PVR) ≥2 Wood Units (WU) already present with early pulmonary vascular disease and impaired survival. Furthermore, PVR has presented as independent prognostic parameter.

How might this impact on clinical practice or future developments?

► The PVR threshold of ≥3 WU may prevent from an early diagnosis in this patient cohort and should be carefully reconsidered.

INTRODUCTION

Pulmonary hypertension (PH) often aggravates systemic sclerosis (SSc) with negative consequences on exercise capacity, quality of life and survival.[1 2] SSc patients may develop PH due to left heart or lung disease[3] or precapillary PH as pulmonary arterial hypertension (PAH) during the course of the disease, which is the main cause of mortality.[4]

Itinerair Scleroderma and evidence-based detection of pulmonary arterial hypertension in systemic sclerosis (DETECT) studies enabled an early diagnosis of PAH among SSc patients (SSc-APAH).[5 6] In the DETECT study, most of the newly diagnosed SSc-APAH patients presented with only slightly elevated mean pulmonary arterial pressure (mPAP), normal or near-normal mean cardiac output (CO) at rest, slightly elevated right atrial (RA) size and a pulmonary vascular resistance (PVR) values <3 Wood Units (WU).[5] Early diagnosis of SSc-APAH is of utmost importance, since it leads to significant improvement of survival rates through the implementation of PAH therapies, as demonstrated in the Itinerair Scleroderma cohort.[7]

Therefore, a new haemodynamic definition of PAH was proposed at the 6th World Symposium of PH,[8] which lowered the cut-off for mPAP from ≥25 mm Hg (stated in the actual PH guidelines)[9] to >20 mm Hg in combination with pulmonary arterial wedge

pressure (PAWP) ≤15 mm Hg and PVR ≥3 WU. The change in the haemodynamic definition of precapillary PH represents a step towards the upper limit of physiological haemodynamic thresholds as shown in a systematic review of Kovacs *et al*[10] including 1187 healthy subjects presenting with an upper limit of normal for mPAP of 20.6 mm Hg (mean+2 SDs). Subsequent studies performed mainly in SSc patients reported that patients with mildly elevated mPAP (21–24 mm Hg) had reduced exercise capacity, impaired quality of life, decreased right ventricular (RV) output reserve, abnormal pulmonary arterial compliance (PAC) and survival.[11–13] Furthermore, it has been shown that SSc patients with exercise PH assessed by right heart catheterisation (RHC) had a similarly impaired survival as SSc patients with manifest resting PAH.[14]

In a recent study reanalysing SSc patients assessed by RHC, the updated definition did not have a significant impact on reclassification, with only 5% of patients being reclassified as PAH.[15] However, the authors of this study suggested that the PVR criterion ≥3 WU is too conservative since a larger proportion of their patients with mild PH (mPAP of 21–24 mm Hg) had PVR <3 WU. A systematic review by Kovacs *et al*[10 16] supports this assertion showing an upper limit of normal of PVR <1.5 WU throughout all age categories. The addition of 2 SD to this upper limit of normal leads to a PVR threshold of 2 WU. The arbitrarily high cut-off of PVR ≥3 WU was consensus during the 6th World Symposium on PH (WSPH) meeting in order to avoid misclassification of PAH.[8]

The aim of the current study was to analyse the impact of the new haemodynamic definition of precapillary PH in a large cohort of SSc patients referred to two reference PH centres. Clinical characteristics of SSc patients with mPAP 21–24 mm Hg and PVR ≥2 WU were compared with patients with manifest PAH and those with normal haemodynamic values. Furthermore, this study investigates for the first time survival of patients according to PVR (threshold 2 WU) and mPAP.

METHODS

Study population

Consecutive SSc patients with diffuse cutaneous SSc (dc-SSc) or limited cutaneous SSc (lc-SSc)[17] fulfilling the classification criteria of the American College of Rheumatology/European League against Rheumatism[18] were prospectively enrolled within the scope of an early detection of PH programme as previously described.[5 13 15] The patients were referred to two PH centres (Centre for Pulmonary Hypertension of the Thoraxklinik at Heidelberg University Hospital, Germany, and Departments of Rheumatology and Pulmonology, University Hospital Zurich, Switzerland). A part of this cohort has already been analysed and published before.[5 13 15]

The referring specialists were rheumatologists, cardiologists, pulmonologists and general practitioners. Individuals were excluded, if they had an already diagnosed PH via RHC prior to enrolment, an ongoing treatment with PAH drugs, renal insufficiency, systemic arterial hypertension with blood pressure values >180/95 mm Hg at rest or >230/120 mm Hg during exercise despite optimised medical treatment, previous evidence of clinically relevant left heart disease, significant lung disease or if they were pregnant.

All patients underwent a detailed clinical work-up, echocardiography and RHC (for details see online supplementary material).[19]

Study design

In order to evaluate the impact of the new PH definition, patients were divided into three groups according to their resting mPAP values: normal mPAP (mPAP ≤20 mm Hg), mildly elevated mPAP (21–24 mm Hg) and manifest PH (mPAP ≥25 mm Hg). Each group was further divided according to PVR. The frequency of patients with mPAP 21–24 mm Hg as well as mPAP ≥25 mm Hg was analysed using different PVR cut-off values (2, 2.5 and 3 WU). Patients with significant lung disease (VCmax <70% and/or signs of significant interstitial lung disease in CT of the lungs) or significant left heart disease (PAWP >15 mm Hg) were reported in the respective haemodynamic subgroups and were excluded from further comparative analysis. Clinical characteristics, 6 min walking distance (6MWD), RA and RV area, tricuspid annular plane systolic excursion (TAPSE), PAC and N-terminal pro-brain natriuretic peptide (NT-proBNP) were compared between patients with mPAP 21–24 mm Hg and PVR ≥2 WU, SSc patients with normal haemodynamics (mPAP ≤20 mm Hg and PVR <2 WU) and patients with manifest PAH according to the current haemodynamic definition (mPAP ≥25 mm Hg and PVR ≥3 WU). Finally, we compared the survival rates of patients according to PVR values (threshold 2 WU).

Statistical methods

Statistical analyses were conducted by a statistician (NB). Data are described as mean±SD. The frequency distribution of different haemodynamic subgroups and further frequency data are given as number and per cent, respectively.

Comparisons of clinical parameters between patients with mPAP 21–24 mm Hg and PVR ≥2 WU versus patients with normal haemodynamics (mPAP ≤20 mm Hg and PVR <2 WU) as well as comparison of patients with mPAP 21–24 mm Hg and PVR ≥2 WU versus manifest PAH (mPAP ≥25 mm Hg and PVR ≥3 WU) were made with Wilcoxon-Mann-Whitney test. Clinical parameters included 6MWD, TAPSE, RA and RV area, NT-proBNP and PAC.

Survival analysis with comparison of different subgroups with mPAP (≤20, 21–24, ≥25 mm Hg) and PVR (<2, ≥2 WU) was performed with Kaplan-Meier analysis and age-adjusted Cox regression analysis. Multivariate Cox regression analysis was performed for parameters at baseline, including age, mPAP, PVR, CO, cardiac index, PAWP, stroke volume, PAC, TAPSE, 6MWD, sex and WHO functional class. The RHC assessment date was set as baseline for survival analysis. Death was defined as death due to any cause.

Three risk stratification models, including REVEAL 2.0,[20] COMPERA/Swedish approach[21 22] and the French approach,[23] were applied to the whole cohort and subsets of the cohort with PVR <2 and ≥2 WU.

The frequency of different phenotypes was calculated from all patients who were included into the study and had valid haemodynamic data, including patients with cardiac or pulmonary disease. Comparisons between haemodynamic subgroups and survival analyses comprised of patients without significant lung disease or PAWP >15 mm Hg.

All analyses have been performed using IBM SPSS 25 (SPSS Statistics V.25, IBM Corporation, Somers, New York, USA).

PATIENT AND PUBLIC INVOLVEMENT STATEMENT

This analysis was aimed at investigating the impact of the new haemodynamic definition of PH. The findings of this analysis will be presented to physicians at national and international congresses and to patients during patient organisation meetings (patient organisations for SSc and for PH) in order to raise awareness of the condition and its early diagnosis. We hope that

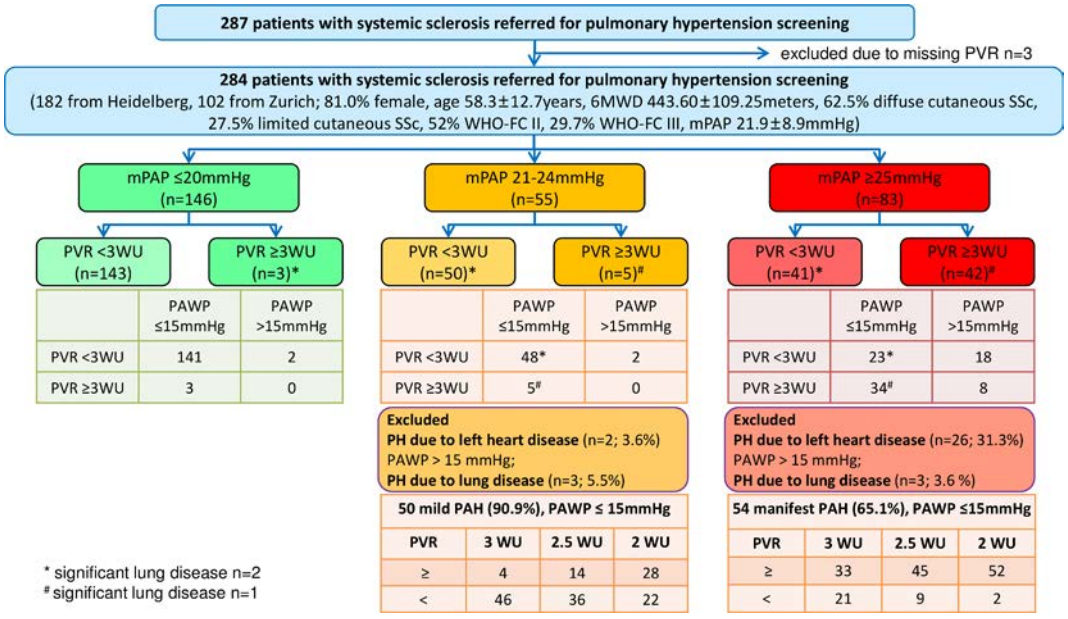


Figure 1 Study flow chart and different PVR thresholds in patients with mPAP 21–24 and ≥25 mm Hg. The flow chart characterises the frequency of haemodynamic subgroups according to mPAP and PVR threshold of 3 WU. The number of patients excluded due to significant lung disease and/or PAWP >15 mm Hg is given for patients with mPAP 21–24 and ≥25 mm Hg. For both haemodynamic groups according to mPAP values, patient numbers are given for different PVR thresholds. According to the new haemodynamic definition, four additional patients with manifest pulmonary arterial hypertension were identified. 6MWD, 6 min walking distance; mPAP, mean pulmonary arterial pressure; PAH, pulmonary arterial hypertension; PAWP, pulmonary arterial wedge pressure; PH, pulmonary hypertension; PVR, pulmonary vascular resistance; SSc, systemic sclerosis; WHO-FC, WHO functional class; WU, Wood Units.

early detection, diagnosis and possibly early treatment of this patient population may be enhanced by our study results.

RESULTS

Baseline characteristics

A total of 287 patients with SSc were screened for PH. Three patients were excluded from the study because of missing haemodynamic data. Thus, the final study group consisted of 284 patients, 182 patients have been assessed in Heidelberg and 102 in Zurich (figure 1, table 1). In the patients included in this study, overt significant lung or left heart disease had been excluded. However, during the assessment at the time of RHC, in 36 out of the 284 patients previously unknown significant lung disease (n=6) or left heart disease with a PAWP >15 mm Hg (n=30) was diagnosed. These 36 patients were excluded from the haemodynamic subgroup analysis, leading to a study cohort without comorbidities of 248 patients (baseline characteristics of analysis set without significant comorbidities, see table 1).

The mean age of the 284 patients of the whole cohort was 58.3±12.7 years, 81.0% were female, 50.7% had dc-SSc and 49.3% lc-SSc. The mean 6MWD was 443.6±109.3 m, 51% were functionally limited with functional class (WHO FC) II, 28.5% had WHO FC III and 0.8% had WHO FC IV (baseline characteristics of analysis set of whole cohort, see table 1).

Patients with manifest PAH (mPAP ≥25 mm Hg, PVR ≥3 WU, PAWP ≤15 mm Hg) were treated with PAH medication according to the guidelines. Patients who did not meet the criteria of manifest PAH were not treated with continuous PAH medication. PAH-targeted therapies including bosentan that is also used for prevention of digital ulcers in SSc were given as non-continuous treatment in less than 5% of patients.

Impact of the new definition of PAH

Out of 284 patients, 146 patients had mPAP ≤20 mm Hg, 55 patients 21–24 mm Hg (mildly elevated mPAP) and 83 patients had mPAP ≥25 mm Hg (figure 1). In the group of mildly elevated mPAP, four patients could be reclassified as manifest SSc-APAH according to the new haemodynamic definition of PAH (1.4% of the whole cohort or 8% among patients with mPAP 21–24 mm Hg; table 2). In these four patients, significant lung or left heart disease had been excluded. The clinical characteristics of these patients are summarised in table 2.

Among the patients in the group with mPAP 21–24 mm Hg, 28 (9.85% of total cohort) had a PVR ≥2 WU with no significant left heart or lung disease and would be newly diagnosed as SSc-APAH using this PVR threshold. In patients with mPAP 21–24 mm Hg and PVR ≥2 WU, 21 out of 25 patients who received exercise RHC could be defined as exercise PH with mPAP >30 mm Hg and total pulmonary resistance >3 WU. No patient had a PAWP >25 mm Hg. Among patients with mPAP ≥25 mm Hg and PAWP ≤15 mm Hg (n=54), 33 had a PVR ≥3 WU, another 19 patients (6.7% of the total cohort) had a PVR ≥2 WU and would be classified as SSc-APAH with a lower PVR threshold. Three out of 146 SSc patients with mPAP ≤20 mm Hg had a PVR ≥3 WU.

Comparison of clinical parameters

The comparisons of clinical parameters have been performed in the cohort without cardiac or pulmonary comorbidities (n=248). Patients with mPAP 21–24 mm Hg and PVR ≥2 WU showed reduced TAPSE (20.6±5.7 vs 23.8±3.9 mm, p=0.004) (table 3) and decreased 6MWD (413.5±99.6 vs 487.7±100.6 m, p<0.001) compared with patients with normal haemodynamics (figure 2).

Table 1 Clinical characteristics of the study cohort with and without concomitant left heart or lung disease

	Whole cohort (n=284)			Cohort without significant cardiac or pulmonary comorbidities (n=248)		
Parameter (unit)	**n**	**Mean±SD**	**95% CI**	**n**	**Mean±SD**	**95% CI**
Female sex, no (%)		238 (82.9)			207 (82.8)	
Age (years)	282	58.28±12.73	56.79 to 59.77	246	57.51±12.69	55.92 to 59.11
Height (cm)	271	165.63±8.41	164.63 to 166.64	238	165.82±8.50	164.74 to 166.91
Weight (kg)	272	70.11±15.5	68.30 to 71.92	240	70.36±15.50	68.38 to 72.33
Systolic blood pressure (mm Hg)	262	130.43±21.50	127.81 to 133.04	234	129.95±20.78	127.28 to 132.63
Diastolic blood pressure (mm Hg)	264	76.03±12.09	74.57 to 77.50	205	75.32±11.41	73.85 to 76.78
WHO FC, no (%)	263			230		
I		52 (19.8)			50 (21.7)	
II		134 (51.0)			121 (52.6)	
II		75 (28.5)			58 (25.2)	
IV		2 (0.8)			1 (0.4)	
SSc subgroups				148		
Diffuse		144 (50.7)			130 (52.4)	
Limited		140 (49.3)			118 (47.6)	
SSc disease duration	276	10.55±15.01	8.77 to 12.34	241	10.70±15.43	8.74 to 12.66
Digital ulcers		79 (29.3%)			70 (29.5)	
Arterial hypertension		92 (34.8%)			78 (33.2)	
Haemodynamics at rest						
mPAP (mm Hg)		21.9±8.89	20.86 to 22.94		20.52±7.87	19.54 to 21.51
PAWP (mm Hg)		10.34±4.55	9.81 to 10.87		9.17±3.14	8.78 to 9.56
CO (L/min)		5.53±1.45	5.36 to 5.69		5.53±1.41	5.35 to 5.71
CI (L/min/m^2)	282	3.14±0.74	3.05 to 3.23	246	3.17±0.75	3.07 to 3.26
PVR (WU)		2.25±1.77	2.05 to 2.46		2.17±1.62	1.97 to 2.37
PAC (mL/mm Hg)	229	5.05±15.70	4.70 to 5.41	229	5.05±2.74	4.70 to 5.41
Echocardiography at rest						
sPAP (mm Hg)	259	32.01±14.05	30.29 to 33.73	225	30.42±13.25	28.68 to 32.16
RA area (cm^2)	231	13.57±4.54	12.98 to 14.15	231	13.57±4.54	12.98 to 14.95
RV area (cm^2)	215	15.35±4.18	14.78 to 15.91	215	15.35±4.18	14.78 to 15.91
TAPSE (mm)	232	22.87±4.46	22.29 to 23.45	232	22.87±4.46	22.29 to 23.45
Lung function						
VCmax (%)	251	92.45±23.12	89.58 to 95.33	219	93.34±22.46	90.35 to 96.33
FEV1 (%)	248	88.06±23.57	85.11 to 91.01	217	88.98±23.50	85.83 to 92.12
TLC (%)	245	93.08±22.85	90.21 to 95.96	215	93.99±22.56	90.95 to 97.02
6MWD						
6MWD (m)	255	443.60±109.25	430.13 to 457.08	227	453.07±105.75	439.24 to 466.90
Laboratory						
NT-proBNP (pg/mL)	205	555.29±1947.85	287.06 to 823.52	205	555.29±1947.85	287.06 to 823.52

In case of missing data, sample sizes are given in brackets.
SSc disease duration defined from first non-Raynaud symptom.
Digital ulcers as defined by Suliman *et al*.[33]

CI, cardiac index; CO, cardiac output; FEV1, forced expiratory volume in first second; mPAP, mean pulmonary arterial pressure; 6MWD, 6 min walking distance; NT-proBNP, N-terminal pro-brain natriuretic peptide; PAC, pulmonary arterial compliance; PAWP, pulmonary arterial wedge pressure; PVR, pulmonary vascular resistance; RA, right atrial; RV, right ventricular; sPAP, systolic pulmonary arterial pressure; SSc, systemic sclerosis; TAPSE, tricuspid annular plane systolic excursion; TLC, total lung capacity; VC, vital capacity; WHO FC, WHO functional class; WU, Wood Units.

PAC significantly differed between patients with mildly elevated mPAP with PVR ≥2 WU and patients with normal haemodynamics (4.02±1.32 vs 6.16±2.84 mL/mm Hg, p<0.001), as well as between mildly elevated mPAP with PVR ≥2 WU and SSc patients with manifest PAH with PVR ≥3 WU (4.02±1.32 vs 2.28±0.99 mL/mm Hg, p<0.001) (figure 2). Stroke volume index (SVI) and diffusing capacity of the lung for carbon monoxide did not significantly differ between groups. RA area significantly differed between patients with mPAP 21–24 mm Hg and PVR ≥2 WU and patients with manifest PAH (p=0.043); RV area showed a difference in trend (p=0.065). NT-proBNP showed comparable results in patients with normal haemodynamics and mPAP 21–24 mm Hg and PVR ≥2 WU. Patients with manifest PAH had significantly higher NT-proBNP than patients with mPAP 21–24 mm Hg and PVR ≥2 WU (p=0.003).

Survival among SSC patients

The survival analyses have been performed in the cohort without comorbidities (n=248) (figure 3). Patients with PVR ≥2 WU showed a significantly worse survival than patients with PVR <2 WU with 1-, 3-, 5- and 7-year survival rate of 97.7%, 90.7%,

Table 2 Detailed characteristics of newly identified patients according to the new haemodynamic definition

										Haemodynamics at rest					Echocardiography at rest					Laboratory
ID	Gender	Age (years)	Height (cm)	Weight (kg)	WHO FC	SSc subgroups	SSc disease duration (years)	Digital ulcers	Arterial hypertension	mPAP (mm Hg)	PAWP (mm Hg)	CO (L/min)	CI (L/min/m²)	PVR (WU)	sPAP (mm Hg)	RA area (cm²)	RV area (cm²)	TAPSE (mm)	6MWD (m)	NT-proBNP (pg/mL)
Patient #1	Male	54	173	104	2	Diffuse	9	No	No	21	3	5.8	6.7	3.1	30	18	28	23	402	N/A
Patient #2	Female	50	165	93	1	Diffuse	10.2	Yes	No	23	10	4.3	2.2	3.0	30	19	27	18	534	854
Patient #3	Female	64	162	51	1	Diffuse	0.4	No	No	24	13	3.4	2.2	3.2	26	9	12	23	606	135
Patient #4	Female	64	162	65	2	Diffuse	1.0	No	No	21	4	5.1	3.1	3.3	35	14	16	31	468	238

CI, cardiac index; CO, cardiac output; mPAP, mean pulmonary arterial pressure ; 6MWD, 6 min walking distance; NT-proBNP, N-terminal pro-brain natriuretic peptide; PAWP, pulmonary arterial wedge pressure; PVR, pulmonary vascular resistance; RA, right atrial; RV, right ventricular; sPAP, systolic pulmonary arterial pressure; SSc, systemic sclerosis; TAPSE, tricuspid annular plane systolic excursion; WHO FC, World Health Organization functional class; WU, Wood Units.

79.4% and 54.3% versus 100%, 94.2%, 91% and 84.2% (Kaplan-Meier p=0.002; age-adjusted Cox regression p=0.028). Survival according to mPAP also significantly differed between groups (Kaplan-Meier p=0.007; age-adjusted Cox regression different in trend p=0.064). In patients with mPAP 21–24 mm Hg, PVR ≥2 WU was identified as significant predictor of survival (Kaplan-Meier p=0.047). Age-adjusted Cox regression showed inconsistent results (p>0.05). In the multivariate analysis at baseline, PVR and 6MWD were independent prognostic parameters for survival.

In patients with PVR <2 WU, 12 out of 147 patients (8.2%) died during follow-up (mean follow-up time 4.2±2.7 years). In patients with PVR ≥2 WU, 18 out of 101 patients (17.8%) died during follow-up (mean follow-up time 3.5±2.6 years). Reasons for death were related to pulmonary vascular disease PAH (3 vs 9 in PVR <2 WU and ≥2 WU, respectively) and not related to pulmonary vascular disease (9 vs 9 in PVR <2 WU and ≥2 WU, respectively). Death was unknown in 4 versus 2 patients with PVR <2 WU and ≥2 WU, respectively.

Three risk stratification models showed significantly different survival for the different risk scores: COMPERA/Swedish approach (p<0.0001)[21 22] and REVEAL 2.0 (p<0.0001),[20] French approach (p=0.001).[23] When stratifying the cohort according to the PVR threshold <2 WU and ≥2 WU, the REVEAL risk score showed significantly different survival for both cohorts (both p<0.0001), whereas COMPERA/Swedish approach did not show significant differences for patients with PVR <2 WU (p=0.35), but for PVR ≥2 WU (p<0.0001) and the French approach did not show significant differences for PVR <2 WU (0.26), but was significant in trend for ≥2 WU (p=0.078).

The application of risk stratification models in patients with mPAP 21–24 mm Hg and PVR 2–3 WU demonstrates that low-risk patients could be detected. The REVEAL risk score was 0–6 (low risk) in 20 out of 28 patients, COMPERA/Swedish approach showed a low risk in 6, a moderate risk for 8 out of 14 patients and the French approach showed a score ≥3/4 (low risk) in 16 out of 28 patients.

DISCUSSION

This is the first large study showing that patients with mild PAH (mPAP 21–24 mm Hg and PVR ≥2 WU, PAWP <15 mm Hg, no significant lung or left heart disease) had already a clinically meaningful pulmonary vascular disease, RV dysfunction and markedly reduced long-term survival. The new haemodynamic definition of PAH as proposed during the 6th WSPH did not have a significant impact on reclassification in the cohort of this study, with only 4 out of 284 patients being newly classified as PAH (1.4% of the total cohort; 8.0% of patients with mild PAH). The number of newly diagnosed patients with mild PAH would however markedly increase to 9.85% in this cohort, if a physiological haemodynamic threshold for PVR (ie, ≥2 WU) would be used. PVR and 6MWD were identified as independent prognostic parameters for survival. Thus, the data of this study show that a PVR threshold ≥3 WU is too high to enable an early diagnosis of PAH and should be replaced by a cut-off value of ≥2 WU.

Impact of the new haemodynamic definition of PH

The results of this bicentric study are in line with data recently published by Jaafar *et al*[15], reporting a small addition to the SSc-APAH diagnosed patients of 4% from the single-centre cohort of the University of Michigan. In addition to the study of Jaafar *et al*, the data of our cohorts document the reduced survival and its association with PVR as independent prognostic predictor. The very useful editorial comment by Kovacs and Olschewski[24] pointed to another important issue that is also addressed in our study. Nineteen out of 50 patients (38%, or 6.7% out of the total cohort) with clearly elevated mPAP (≥25 mm Hg) and no relevant left heart (PAWP <15 mm Hg) or lung disease failed to fulfil the haemodynamic criteria of precapillary PH, because they had a PVR <3 WU but ≥2 WU. As already shown in the DETECT study, many of the patients diagnosed by a systematic screening programme including RHC have a normal CO at rest[5] and therefore rarely present with PVR values ≥3 WU. However, in these patients, we previously detected an RV dysfunction with markedly impaired cardiac reserve (reduced CO during exercise).[13] The strength of our study is the addition of survival data, supporting the hypothesis that a PVR ≥2 WU is already of great prognostic importance for the patients.

Furthermore, the data of this study document for the first time that patients with mPAP 21–24 mm Hg and PVR ≥2 WU already show impaired exercise capacity, right heart systolic function and PAC. SVI did not significantly differ between groups, though this parameter has been shown to be an important prognostic predictor of outcomes.[25 26] Exercise performance and right heart systolic function could therefore already be damaged with an only slight increase of PVR and still normal right heart size, stroke volume and NT-proBNP. Natriuretic peptides are released in response to myocardial stretch and patients with an early stage of the disease usually have normal right heart size.[27] Right heart enlargement is typically a hallmark of advanced disease, more prominent when haemodynamic decompensation occurs.[28]

Thus, in an update of the proposed PAH definition, the PVR threshold should be lowered to ≥2 WU, since only this value would allow an early diagnosis of precapillary PH. The PVR cut-off value ≥3 WU does not allow an early PAH diagnosis and is associated with a markedly reduced long-term survival in patients with SSc and early pulmonary vascular disease.

Clinical implications

Today, there are no sufficient data whether patients with mildly elevated mPAP (21–24 mm Hg) and PVR ≥2 WU should be

Table 3 Characteristics of haemodynamic subgroups

	A) mPAP 21–24 mm Hg, PVR ≥2 WU (n=28)			B) mPAP ≤20 mm Hg, PVR <2 WU 'normal' (n=123)			C) mPAP ≥25 mm Hg, PVR ≥3 WU 'manifest PAH' (n=33)		
Parameter (unit)	**Mean±SD or n (%)**	**95% CI**	**n**	**Mean±SD or n (%)**	**95% CI**	**n**	**Mean±SD or n (%)**	**95% CI**	**n**
Female sex, no (%)	24 (85.7)			107 (85.6)			28 (84.8)		
Age (years)	61.22±8.79	57.57 to 64.70	27	54.32±12.53	52.07 to 56.27	122	62.27±12.24	57.93 to 66.61	
Height (cm)	162.81±9.87	158.91 to 166.72	27	166.97±7.35	165.61 to 168.32	116	162.97±10.49	159.25 to 166.69	
Weight (kg)	72.04±16.18	65.64 to 78.45	27	69.90±14.74	67.21 to 72.59	118	66.22±14.13	61.13 to 71.32	32
Systolic blood pressure (mm Hg)	139.65±21.85	130.65 to 130.83	26	129.69±20.19	123.00 to 130.37	118	126.89±19.71	119.25 to 134.54	28
Diastolic blood pressure (mm Hg)	75.12±8.69	71.60 to 78.63	26	76.29±10.32	74.42 to 78.17	119	73.52±67.4	67.40 to 79.64	29
WHO FC, no (%)									
I	6 (23.0)			34 (29.8)			1 (3.4)		
II	11 (42.3)			65 (57.0)			8 (27.6)		
II	8 (30.7)			15 (13.2)			20 (68.9)		
IV	1 (3.8)			0 (0)			0 (0)		
SSc subgroups									
Diffuse	17 (60.8)			63 (51.2)			13 (39.4)		
Limited	11 (39.2)			59 (48.8)			18 (60.6)		
SSc disease duration	12.98±22.18	4.38 to 21.58		10.17±13.76	7.67 to 12.66	119	9.42±7.32	6.42 to 12.41	32
Digital ulcera	6 (24.0)			36 (30.3)			13 (41.9)		
Arterial hypertension	11 (40.7)			33 (28.4)			12 (40.0)		
Haemodynamics at rest									
mPAP (mm Hg)	22.39±1.03	21.99 to 22.79		15.28±3.11	14.73 to 15.84		35.67±7.27	33.09 to 38.24	
PAWP (mm Hg)	8.82±2.93	7.68 to 9.96		8.54±2.99	8.00 to 9.07		9.39±3.11	8.29 to 10.5	
CO (L/min)	5.47±1.11	5.04 to 5.90		5.62±1.51	5.35 to 5.89		5.02±1.33	4.55 to 5.50	
CI (L/min/m^2)	3.18±0.72	2.9 to 3.46		3.19±0.77	3.05 to 3.33	121	2.95±0.72	2.70 to 3.21	
PVR (WU)	2.51±0.35	2.37 to 2.64		1.24±0.43	1.16 to 1.31		5.47±2.05	4.75 to 6.20	
PAC (mL/mm Hg)	4.02±1.32	3.48 to 4.57	25	6.16±2.84	5.64 to 6.68	117	2.28±0.99	1.89 to 2.66	28
SVI (mm Hg/L*min^{-1})	42.1±9.2	38.3 to 45.9	25	44.0±11.3	41.8 to 46.1	109	40.7±8.2	37.6 to 43.8	30
Echocardiography at rest									
sPAP (mm Hg)	29.42±9.81	25.46 to 33.38	26	24.42±6.19	23.25 to 25.59	110	50.97±19.48	43.69 to 58.24	30
RA area (cm^2)	13.60±4.64	11.65 to 15.56	24	12.75±4.01	12.02 to 13.48	119	16.35±5.78	14.06 to 18.63	27
RV area (cm^2)	16.13±5.68	13.67 to 18.59	23	14.76±3.67	14.06 to 15.46	109	18.03±4.30	16.26 to 19.81	25
TAPSE	20.59±5.72	18.33 to 22.86	27	23.79±3.92	23.07 to 24.51	117	20.02±4.35	18.36 to 21.67	29
6MWD									
6MWD (m)	413.54±99.63	373.30 to 453.78	26	487.74±100.60	469.24 to 506.24	116	383.35±109.03	339.31 to 427.38	23
Laboratory									
NT-proBNP (pg/mL)	354.56±505.87	145.75 to 563.37	25	324.30±726.94	179.32 to 469.29	99	1091.14±1647.03	452.49 to 1729.79	28
Lung function									
DLCO (%)	53.9±18.6	46.7 to 61.1	28	61.5±13.2	59.1 to 64.0	119	43.7±17.0	37.3 to 50.1	30

In case of missing data, sample sizes are given in brackets.
*P values refer to Wilcoxon-Mann-Whitney test.
CI, cardiac index; CO, cardiac output; DLCO, diffusing capacity of the lung for carbon monoxide; mPAP, mean pulmonary arterial pressure;6MWD, 6 min walking distance; NT-proBNP, N-terminal pro-brain natriuretic peptide; PAC, pulmonary arterial compliance;PAH, pulmonary arterial hypertension; PAWP, pulmonary arterial wedge pressure; PVR, pulmonary vascular resistance; RA, right atrial; RV, right ventricular; sPAP, systolic pulmonary arterial pressure;SSc, systemic sclerosis; SVI, stroke volume index; TAPSE, tricuspid annular plane systolic excursion; WHO WHO FC, WHO functional class; WU, Wood Units.

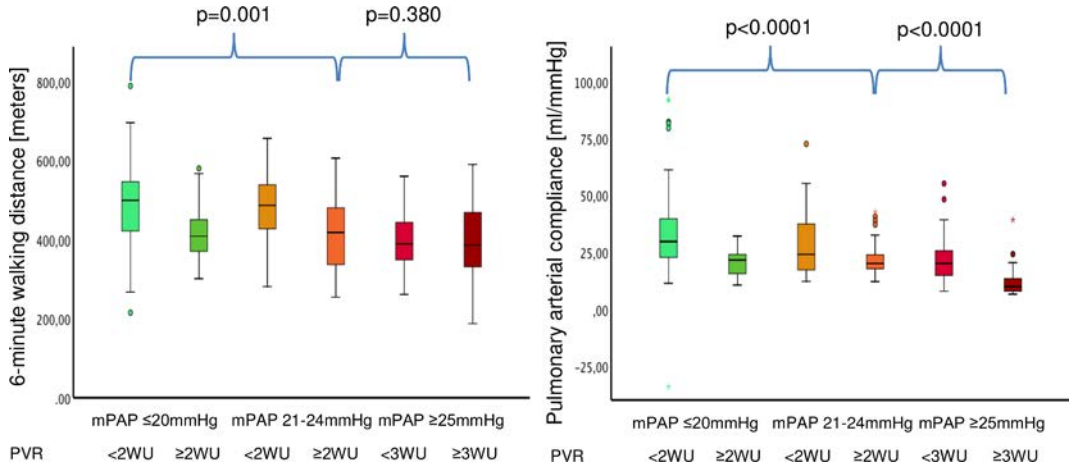


Figure 2 Six-minute walking distance (6MWD) and pulmonary arterial compliance (PAC) in different haemodynamic subgroups. Patients with mPAP 21–24 mm Hg and PVR ≥2 WU showed a significantly lower 6MWD than patients with mPAP ≤20 mm Hg (t-test p=0.001), but did not significantly differ from patients with mPAP ≥25 mm Hg. Patients with mPAP 21–24 mm Hg and PVR ≥2 WU also showed a significantly lower PAC than patients with mPAP ≤20 mm Hg (t-test p<0.0001), and a significantly higher PAC than patients with mPAP ≥25 mm Hg (t-test p<0.0001). The bracket ends in the graph point to the two groups that were compared by Student's t-test. mPAP, mean pulmonary arterial pressure; PVR, pulmonary vascular resistance; WU, Wood Units.

treated with PAH medication. Two open-labelled pilot studies investigating the effect of endothelin receptors antagonists showed promising results.[29 30] A recent double-blind randomised controlled trial early treatment of borderline pulmonary arterial hypertension associated with SSc (EDITA) testing the effect of ambrisentan among SSc patients with mildly elevated mPAP and/or exercise PH failed to change the primary endpoint change of mPAP at rest but showed an improvement of PVR as secondary endpoint,[31] which may be of prognostic relevance in this patient cohort. Further research in this field is needed.

Two of three risk stratification models only showed significant differences between groups in patients with PVR ≥2 WU. Thus, our data show that the application of different risk models to our cohort is mainly applicable in patients with early pulmonary vascular disease (PVR ≥2 WU) and not in patients with normal haemodynamics (<2 WU).

Strengths and study limitations

This study provides important insights for the diagnosis of early SSc-APAH. All patients were screened in experienced centres and had a thorough clinical assessment including RHC and were followed up taking their comorbidities (heart or lung diseases) into account. RHC follow-up data to describe the subsequent clinical course of the patients would have been desirable and should be aimed for with future studies.

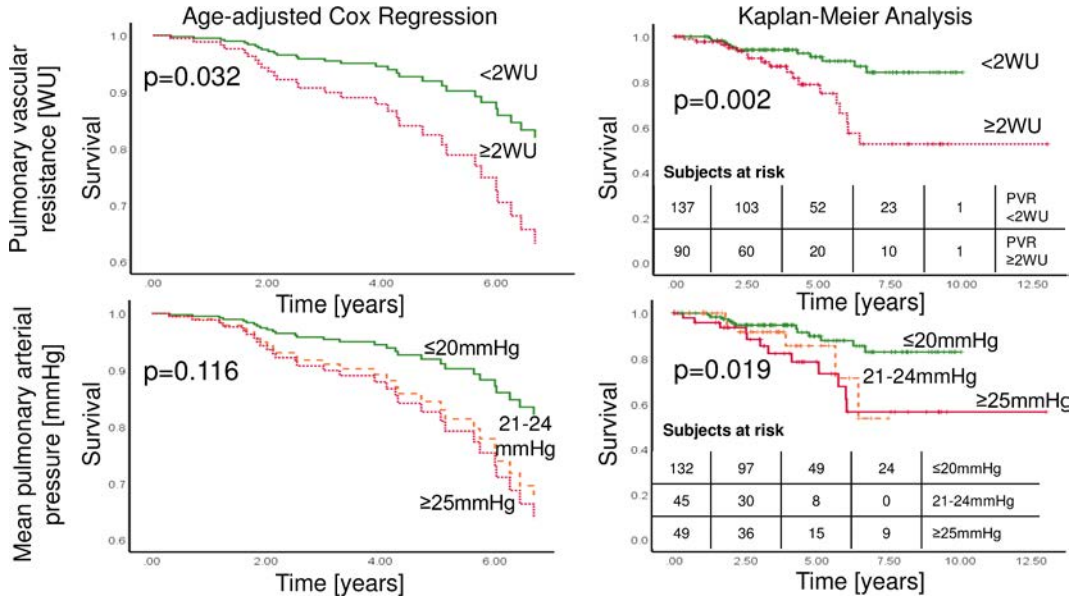


Figure 3 Survival analysis in different haemodynamic groups according to PVR and mPAP. Age-adjusted Cox regression is shown on the left hand, Kaplan Meier analysis on the right hand. Patients with PVR <2 WU showed a significantly better survival than patients with PVR ≥2 WU with 1-, 3-, 5- and 7-year survival rate of 100%, 94.2%, 91% and 84.2% versus 97.7%, 90.7%, 79.4% and 54.3%. Survival of haemodynamic subgroups according to mPAP significantly differed between groups (Kaplan-Meier p=0.007; age-adjusted Cox regression different in trend p=0.064) (figure 7). Patients with significant left heart or lung disease were excluded from the analysis. mPAP, mean pulmonary arterial pressure; PVR, pulmonary vascular resistance; WU, Wood Units

The highest proportion of patients with dc-SSc could be found in the group of patients with mPAP 21–24 mm Hg and PVR ≥2 WU (60.8%), whereas the lowest proportion (39.4%) was in the group with mPAP ≥25 mm Hg and PVR ≥3 WU. The higher proportion of dc-SSc may well have influenced 6MWD, as these patients show lower exercise capacity.[32] However, as dc-SSc has shown to have less severe alteration of haemodynamics,[32] our findings are congruent with early pulmonary vascular disease.

CONCLUSION

The data of this study show that a PVR threshold ≥3 WU is likely too high to enable an early diagnosis of PAH. In the group of mildly elevated mPAP, only four patients (1.4% of the whole SSc-cohort) had PVR values ≥3 WU and could be reclassified as manifest SSc-APAH according to the new definition. A PVR threshold ≥2 WU was already associated with pulmonary vascular disease and significantly reduced survival and would be more appropriate for the early diagnosis of SSc-APAH among this high-risk population. Further studies are needed to analyse the impact of the new PAH definition in other risk groups and to investigate whether patients with mildly elevated mPAP (21–24 mm Hg) and PVR ≥2 WU should be treated with PAH medication.

Author affiliations
[1]Centre for Pulmonary Hypertension, Thoraxklinik Heidelberg gGmbH at Heidelberg University Hospital, Heidelberg, Germany
[2]Translational Lung Research Center Heidelberg (TLRC), Member of the German Center for Lung Research (DZL), Heidelberg, Germany
[3]Department of Rheumatology, University Hospital Zurich, Zurich, Switzerland
[4]IRCCS SDN Naples, Naples, Italy
[5]Department of Internal Medicine V: Hematology, Oncology and Rheumatology, University of Heidelberg, Heidelberg, Germany
[6]Department of Pediatric Cardiology, University Hospital Heidelberg, Heidelberg, Germany
[7]Lung centre, Klinikum Mittelbaden gGmbH, Baden-Baden Balg, Germany
[8]Department of Pulmonology, University Hospital Zurich, Zurich, Switzerland

Acknowledgements We would like to thank all patients who participated in the study. This work was the doctoral thesis of NM.

Contributors PX, SJ and NM made substantial contributions to the acquisition, analysis and interpretation of data, drafting and revision of the work. This is the doctorate thesis of NM. VT, ML and AMM made substantial contributions to the acquisition and interpretation of data and drafting of the work. NBlank made substantial contributions to the conception, acquisition and interpretation of data, and revision of the work. BE, MG, SH, CN, CB and SU made substantial contributions to the acquisition and interpretation of data, and revision of the work. H-ML and EG made substantial contributions to the conception, acquisition, analysis and interpretation of data, revision of the work. NBenjamin made substantial contributions to the conception, acquisition, data analysis and interpretation of data, drafting and revision of the work. OD made substantial contributions to the conception, acquisition, analysis and interpretation of data, drafting and revision of the work. All authors read and approved the manuscript and agree to all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests PX has received personal fees from MSD and OMT. AM has received personal fees from Bayer, outside the submitted work. NBlank has received consulting fees, speaker fees and/or honoraria from MSD, GSK, Actelion and Bayer Vital. BE received travel fees, consulting fees, speaking fees and/or honoraria from Actelion, MSD, Bayer and OMT (less than $10 000 each). MG serves on an advisory Board for Bayer AG and receives lecture fees from Pfizer, MSD and Actelion Pharmaceuticals. SH has received personal fees from Bayer, MSD, Actelion and GSK, outside the submitted work. H-ML has received consulting fees, speaking fees and/or honoraria from AbbVie, BMS, Pfizer, Cellgene, Medac, GSK, Roche, Chugai, Novartis, UCB, Janssen-Cilag, AstraZeneca and Lilly (less than $10 000 each) and research support from AbbVie, MSD, BMS, Cellgene, Medac, GSK, Roche, Chugai, Novartis, UCB, Janssen-Cilag, AstraZeneca, Lilly, Baxter, SOBI, Biogen, Actelion, Bayer Vital, Shire, Octapharm, Sanofi, Hexal, Mundipharm and Thermo Fisher. CN has received consulting fees, speaking fees and/or honoraria from Actelion, MSD, Boehringer, Novartis, Bayer and AstraZeneca (less than $10 000 each). CB received travel support from Actelion SA and Orpha Swiss and speakers' fee from MSD. SU received grants from the Swiss National Science foundation and the Zurich Lung League, travel support and speakers' fees from Actelion SA, Switzerland, Bayer SA, Germany, MSD, Switzerland and Orpha Swiss. EG has received grants and personal fees from Actelion, Bayer AG and MSD; grants from GSK, Novartis, and United Therapeutics; and personal fees from SCOPE, OrPha Swiss GmbH, and Zurich Heart House (less than $10 000 each). NBenjamin received speaker fees from Actelion pharmaceuticals, Bayer HealthCare and MSD. OD has received grants and personal fees from research consultancies from AM, Acceleron Pharma, Amgen, AnaMar, Bayer, Beacon Discovery, Boehringer Ingelheim, Catenion, CSL Behring, ChemomAb, Ergonex, GSK, Inventiva, Italfarmaco, iQone, iQvia, Lilly, medac, Medscape, Mitsubishi Tanabe Pharma, MSD, Novartis, Pfizer, Roche, Sanofi, Blade Therapeutics, Glenmark Pharmaceuticals, Target Bio Science and UCB, outside the submitted work, to investigate potential treatments of scleroderma and its complications. In addition, OD has a patent mir-29 for the treatment of systemic sclerosis issued (US8247389, EP2331143).

Patient consent for publication Not required.

Ethics approval The ethics committees of the Medical Faculty Heidelberg (S360/2009 and S231/2019) and Zurich (EUSTAR—database 839/BASEC Nr. 2016-01515 and cohort study BASEC-Nr. 2018-02165) had no objection against the conduct of the trial.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

ORCID iDs
Nicola Benjamin http://orcid.org/0000-0003-4730-969X
Oliver Distler http://orcid.org/0000-0002-0546-8310

REFERENCES

1 Condliffe R, Kiely DG, Peacock AJ, *et al*. Connective tissue disease-associated pulmonary arterial hypertension in the modern treatment era. *Am J Respir Crit Care Med* 2009;179:151–7.
2 Hao Y, Hudson M, Baron M, *et al*. Early mortality in a multinational systemic sclerosis inception cohort. *Arthritis Rheumatol* 2017;69:1067–77.
3 Lefèvre G, Dauchet L, Hachulla E, *et al*. Survival and prognostic factors in systemic sclerosis-associated pulmonary hypertension: a systematic review and meta-analysis. *Arthritis Rheum* 2013;65:2412–23.
4 Koh E*et al*. Pulmonary hypertension in systemic sclerosis: an analysis of 17 patients. *Rheumatology* 1996;35:989–93.
5 Coghlan JG, Denton CP, Grünig E, *et al*. Evidence-based detection of pulmonary arterial hypertension in systemic sclerosis: the detect study. *Ann Rheum Dis* 2014;73:1340–9.
6 Hachulla E, Gressin V, Guillevin L, *et al*. Early detection of pulmonary arterial hypertension in systemic sclerosis: a French nationwide prospective multicenter study. *Arthritis Rheum* 2005;52:3792–800.
7 Humbert M, Yaici A, de Groote P, *et al*. Screening for pulmonary arterial hypertension in patients with systemic sclerosis: clinical characteristics at diagnosis and long-term survival. *Arthritis Rheum* 2011;63:3522–30.
8 Simonneau G, Montani D, Celermajer DS, *et al*. Haemodynamic definitions and updated clinical classification of pulmonary hypertension. *Eur Respir J* 2019;53.
9 Galiè N, Humbert M, Vachiery J-L, *et al*. 2015 ESC/ERS guidelines for the diagnosis and treatment of pulmonary hypertension: the joint Task force for the diagnosis and treatment of pulmonary hypertension of the European Society of cardiology (ESC) and the European respiratory Society (ERS): endorsed by: association for European paediatric and congenital cardiology (AEPC), International Society for heart and lung transplantation (ISHLT). *Eur Heart J* 2016;37:67–119.
10 Kovacs G, Berghold A, Scheidl S, *et al*. Pulmonary arterial pressure during rest and exercise in healthy subjects: a systematic review. *Eur Respir J* 2009;34:888–94.
11 Valerio CJ, Schreiber BE, Handler CE, *et al*. Borderline mean pulmonary artery pressure in patients with systemic sclerosis: transpulmonary gradient predicts risk of developing pulmonary hypertension. *Arthritis Rheum* 2013;65:1074–84.
12 Coghlan JG, Wolf M, Distler O, *et al*. Incidence of pulmonary hypertension and determining factors in patients with systemic sclerosis. *Eur Respir J* 2018;51.
13 Nagel C, Marra AM, Benjamin N, *et al*. Reduced right ventricular output reserve in patients with systemic sclerosis and mildly elevated pulmonary artery pressure. *Arthritis Rheumatol* 2019;71:805–16.
14 Stamm A, Saxer S, Lichtblau M, *et al*. Exercise pulmonary haemodynamics predict outcome in patients with systemic sclerosis. *Eur Respir J* 2016;48:1658–67.
15 Jaafar S, Visovatti S, Young A, *et al*. Impact of the revised haemodynamic definition on the diagnosis of pulmonary hypertension in patients with systemic sclerosis. *Eur Respir J* 2019;54:1900586.
16 Kovacs G, Olschewski A, Berghold A, *et al*. Pulmonary vascular resistances during exercise in normal subjects: a systematic review. *Eur Respir J* 2012;39:319–28.
17 LeRoy EC, Medsger TA. Criteria for the classification of early systemic sclerosis. *J Rheumatol* 2001;28:1573–6.

18 van den Hoogen F, Khanna D, Fransen J, *et al*. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League against rheumatism collaborative initiative. *Arthritis Rheum* 2013;65:2737–47.
19 Kovacs G, Avian A, Olschewski A, *et al*. Zero reference level for right heart catheterisation. *Eur Respir J* 2013;42:1586–94.
20 Benza RL, Gomberg-Maitland M, Elliott CG, *et al*. Predicting survival in patients with pulmonary arterial hypertension: the reveal risk score calculator 2.0 and comparison with ESC/ERS-Based risk assessment strategies. *Chest* 2019;156:323–37.
21 Hoeper MM, Pittrow D, Opitz C, *et al*. Risk assessment in pulmonary arterial hypertension. *Eur Respir J* 2018;51.
22 Kylhammar D, Kjellström B, Hjalmarsson C, *et al*. A comprehensive risk stratification at early follow-up determines prognosis in pulmonary arterial hypertension. *Eur Heart J* 2018;39:4175–81.
23 Boucly A, Weatherald J, Humbert M, *et al*. Risk assessment in pulmonary arterial hypertension. *Eur Respir J* 2018;51.
24 Kovacs G, Olschewski H. Debating the new haemodynamic definition of pulmonary hypertension: much ado about nothing? *Eur Respir J* 2019;54.
25 Weatherald J, Boucly A, Chemla D, *et al*. Prognostic value of follow-up hemodynamic variables after initial management in pulmonary arterial hypertension. *Circulation* 2018;137:693–704.
26 Weatherald J, Boucly A, Launay D, *et al*. Haemodynamics and serial risk assessment in systemic sclerosis associated pulmonary arterial hypertension. *Eur Respir J* 2018;52.
27 Salzano A, Marra AM, Proietti M, *et al*. Biomarkers in heart failure and associated diseases. *Dis Markers* 2019;2019:8768624.
28 Ferrara F, Gargani L, Ostenfeld E, *et al*. Imaging the right heart pulmonary circulation unit: insights from advanced ultrasound techniques. *Echocardiography* 2017;34:1216–31.
29 Kovacs G, Maier R, Aberer E, *et al*. Pulmonary arterial hypertension therapy may be safe and effective in patients with systemic sclerosis and borderline pulmonary artery pressure. *Arthritis Rheum* 2012;64:1257–62.
30 Saggar R, Khanna D, Shapiro S, *et al*. Brief report: effect of ambrisentan treatment on exercise-induced pulmonary hypertension in systemic sclerosis: a prospective single-center, open-label pilot study. *Arthritis Rheum* 2012;64:4072–7.
31 Pan Z, Marra AM, Benjamin N, *et al*. Early treatment with ambrisentan of mildly elevated mean pulmonary arterial pressure associated with systemic sclerosis: a randomized, controlled, double-blind, parallel group study (EDITA study). *Arthritis Res Ther* 2019;21.
32 Mathai SC, Hassoun PM. Pulmonary arterial hypertension associated with systemic sclerosis. *Expert Rev Respir Med* 2011;5:267–79.
33 Suliman YA, Bruni C, Johnson SR, *et al*. Defining skin ulcers in systemic sclerosis: systematic literature review and proposed world scleroderma Foundation (WSF) definition. *J Scleroderma Relat Disord* 2017;2:115–20.

TRANSLATIONAL SCIENCE

Global skin gene expression analysis of early diffuse cutaneous systemic sclerosis shows a prominent innate and adaptive inflammatory profile

Brian Skaug ,[1] Dinesh Khanna ,[2] William R Swindell,[3,4] Monique E Hinchcliff,[5] Tracy M Frech,[6] Virginia D Steen,[7] Faye N Hant,[8] Jessica K Gordon,[9] Ami A Shah,[10] Lisha Zhu,[11] W Jim Zheng,[11] Jeffrey L Browning,[12] Alexander M S Barron,[12] Minghua Wu ,[1] Sudha Visvanathan ,[13] Patrick Baum,[14] Jennifer M Franks ,[15,16] Michael L Whitfield,[15,16] Victoria K Shanmugam,[17] Robyn T Domsic,[18] Flavia V Castelino,[19] Elana J Bernstein,[20] Nancy Wareing,[1] Marka A Lyons,[1] Jun Ying,[1] Julio Charles,[1] Maureen D Mayes,[1] Shervin Assassi [1]

Handling editor Josef S Smolen

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-215894).

For numbered affiliations see end of article.

Correspondence to
Dr Brian Skaug, Department of Internal Medicine, Division of Rheumatology and Clinical Immunogenetics, University of Texas Health Science Center at Houston, Houston, TX 77030, USA;
brian.a.skaug@uth.tmc.edu

Received 14 June 2019
Revised 7 November 2019
Accepted 7 November 2019
Published Online First 25 November 2019

To cite: Skaug B, Khanna D, Swindell WR, *et al*. *Ann Rheum Dis* 2020;**79**:379–386.

ABSTRACT

Objectives Determine global skin transcriptome patterns of early diffuse systemic sclerosis (SSc) and how they differ from later disease.

Methods Skin biopsy RNA from 48 patients in the Prospective Registry for Early Systemic Sclerosis (PRESS) cohort (mean disease duration 1.3 years) and 33 matched healthy controls was examined by next-generation RNA sequencing. Data were analysed for cell type-specific signatures and compared with similarly obtained data from 55 previously biopsied patients in Genetics versus Environment in Scleroderma Outcomes Study cohort with longer disease duration (mean 7.4 years) and their matched controls. Correlations with histological features and clinical course were also evaluated.

Results SSc patients in PRESS had a high prevalence of M2 (96%) and M1 (94%) macrophage and CD8 T cell (65%), CD4 T cell (60%) and B cell (69%) signatures. Immunohistochemical staining of immune cell markers correlated with the gene expression-based immune cell signatures. The prevalence of immune cell signatures in early diffuse SSc patients was higher than in patients with longer disease duration. In the multivariable model, adaptive immune cell signatures were significantly associated with shorter disease duration, while fibroblast and macrophage cell type signatures were associated with higher modified Rodnan Skin Score (mRSS). Immune cell signatures also correlated with skin thickness progression rate prior to biopsy, but did not predict subsequent mRSS progression.

Conclusions Skin in early diffuse SSc has prominent innate and adaptive immune cell signatures. As a prominently affected end organ, these signatures reflect the preceding rate of disease progression. These findings could have implications in understanding SSc pathogenesis and clinical trial design.

Key messages

What is already known about this subject?

- Skin gene expression is altered in patients with systemic sclerosis (SSc) based on data from microarrays, but heterogeneity exists in skin gene expression profiles of SSc patients.

What does this study add?

- A large-scale analysis of skin transcript expression specifically in patients with early, diffuse cutaneous SSc and comparison to patients with later disease revealed that innate and adaptive immune cell gene expression is more prominent in early diffuse SSc compared with later disease. After adjustment for key clinical characteristics, adaptive immune cell signatures were associated with shorter disease duration.
- Immune cell signatures appeared to reflect preceding skin thickness progression rate but did not predict subsequent modified Rodnan Skin Score progression.

How might this impact on clinical practice or future developments?

- The prominence of innate and adaptive immune cell signatures in early diffuse SSc would seem to support the premise of using immune-modulatory therapies in this subgroup of patients.
- There appear to be limitations in the use of skin gene expression profiles to predict subsequent disease progression, perhaps related to heterogeneity among SSc patient cohorts.

INTRODUCTION

Systemic sclerosis (SSc) is a multi-system autoimmune and fibrotic disease associated with high morbidity and mortality.[1 2] Treatment options remain limited, and management is complicated by heterogeneity in clinical course and treatment response.

Whole transcriptome gene expression profiling can yield insights into disease pathogenesis and identify distinct subgroups of patients.[3 4] We and others have previously used microarray technology to measure global gene expression in skin biopsies

from SSc patients in comparison to healthy controls (HCs),[5–12] revealing distinct gene expression patterns in SSc skin. Fibrotic and inflammatory gene expression signatures have been observed in a large percentage of patients, while a subset of patients has 'normal-like' gene expression profiles. These studies highlight heterogeneity in SSc skin gene expression. A large-scale study to characterise skin gene expression specifically in early, diffuse SSc in comparison to those with later stage disease has been lacking.

We investigated the transcript expression profiles of skin specimens from a large group of patients with early, diffuse SSc from the Prospective Registry for Early Systemic Sclerosis (PRESS) cohort using next generation RNA sequencing. These data were compared with HC skin and to patients in the Genetics versus Environment in Scleroderma Outcomes Study (GENISOS), in which patients had a longer average disease duration.

METHODS

Patients and control subjects

Patients were recruited from PRESS, an observational cohort of early diffuse SSc patients from 11 US academic medical centres.[13] Skin biopsies from 48 patients within 3 years of onset of first non-Raynaud's symptom were used for RNA sequencing, along with 33 biopsies from HCs matched to patients by age, sex and ethnicity. Ten repeat biopsies from eight SSc patients were also available. Skin biopsy was optional in PRESS, and all available biopsies in the PRESS cohort at the time of study were included. Patients fulfilled the 2013 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) classification criteria for SSc and had diffuse skin involvement.[14] Modified Rodnan Skin Score (mRSS) and local skin score at the biopsy site were recorded at the time of biopsy. Skin thickness progression rate (STPR) was calculated similarly to what was previously described,[15] using the equation mRSS at the time of biopsy/time from first puffy fingers or skin thickening. Participants provided informed and written consent.

Skin biopsy and RNA sequencing

Punch biopsies were obtained from the forearm skin. The methods for RNA sequencing and analysis are described in online supplementary methods. Data from the PRESS cohort were compared with similarly obtained data from the GENISOS cohort that included SSc patients with longer disease duration at the time of biopsy.[10] Although microarray technology was used for gene expression profiling in the previously published study, we performed RNA sequencing in these GENISOS samples (n=55) and matched HCs (n=33) for the present study in order to avoid heterogeneity resulting from methodological differences.

Analysis of cell type-specific expression

We performed cell type-specific gene expression analysis using the method we have used previously.[10 16] Details are provided in the online supplementary methods.

Assignment of patients to 'intrinsic subsets' based on skin gene expression

Fragments per kilobase million (FPKM) values were sent to JMF and MLW who were blinded to all clinical data and assigned each sample to one of four 'intrinsic subsets' using previously described methods.[17 18]

Immunohistochemistry

Immunohistochemical (IHC) analyses of skin biopsies are described in the online supplementary methods.

Statistical analysis

Associations between cell type signatures and clinical or histological features were analysed by Spearman's rank order correlation. Cell type signature scores were log-transformed and compared between the PRESS and GENISOS cohorts by Student's t-test. Multivariable regression analyses were performed with pooled data from both cohorts with adjustment for clinical variables noted in the text. mRSS and STPR within the intrinsic subsets were analysed by linear regression analyses, using the normal-like subset as a reference.

RESULTS

Demographics

Demographics and clinical characteristics of participants from the PRESS cohort and matched HCs, along with the GENISOS cohort and their matched HCs, are shown in table 1.

Transcript expression profile of early diffuse SSc skin

Three thousand eighty seven transcripts were differentially expressed between SSc patients and HCs using a false discovery rate cut-off of 0.05 and fold change cut-off of >1.5 or <0.67, including 927 long non-coding RNAs (lncRNAs). Unsupervised hierarchical clustering revealed nearly complete discrimination between differential transcript expression in HCs and SSc patients, with the exception of three SSc patients whose transcript expression profile largely resembled that of HCs (figure 1A). Lists of differentially expressed transcripts between SSc and HC and associations between transcripts and mRSS or forced vital capacity (FVC) in SSc patients at the time of skin biopsy are included in the supplementary data file on our webpage (https://www.uth.tmc.edu/scleroderma/).The most over-represented pathways in SSc skin based on Ingenuity Pathway Analysis were hepatic fibrosis, granulocyte and agranulocyte adhesion and diapedesis, and Th1 and Th2 activation pathways (figure 1B). Th1 and Th2 activation pathways had not been previously observed in the skin of SSc patients.[10] The top activated transcriptional regulators were predicted to be signal transducer and activator of transcription 1, interferon regulatory factor 7, and CCAAT enhancer binding protein beta, while the top activated cytokines/growth factors were interferon gamma, tumour necrosis factor and interleukin 1 beta (IL-1β) (figure 1C, D). Surprisingly, transforming growth factor beta (TGFβ) ranked 15th among upstream cytokines/growth factors (data not shown), in contrast to our prior study of patients with longer disease duration in which it had ranked first.[10]

Prominent innate and adaptive immune cell signatures in early diffuse SSc skin

Cell type-specific analysis revealed that most patients had increased innate and adaptive immune cell signatures compared with HCs (figure 2A). The most prevalent signatures upregulated in SSc compared with HC were those of M2 and M1 macrophages (96% and 94% of SSc patients, respectively). A fibroblast signature was present in 92% of patients. Most SSc patients also had CD4 T cell, CD8 T cell and B cell signatures (60%, 65% and 69%, respectively). No significant differences in cell type signatures were observed in male versus female patients or in RNA polymerase III antibody-positive versus topoisomerase I antibody-positive patients (online supplementary tables 1 and 2, respectively).

We compared the cell type signatures in PRESS patients to those of GENISOS patients for whom we had previously performed skin biopsies and analysed RNA expression by microarray.[10] To allow for comparison between the two cohorts, RNA sequencing

Table 1 Demographics and clinical characteristics of SSc patients from PRESS and GENISOS cohorts and matched healthy controls at the time of skin biopsy

Characteristic	PRESS SSc (n=48)	HC matched to PRESS (n=33)	GENISOS SSc (n=55)	HC matched to GENISOS (n=33)
Age (years) at biopsy, mean (SD)	48.0 (15.0)	47.4 (13.4)	52.8 (12.6)	46.8 (11.7)
Race/ethnicity, n (%)				
White	30 (62.5)	24 (72.7)	38 (69.1)	19 (57.6)
Black	5 (10.4)	4 (12.1)	8 (14.5)	8 (24.2)
Hispanic	9 (18.8)	4 (12.1)	9 (16.4)	6 (18.2)
Other	4 (8.3)	1 (3.0)	0	0
Female, n (%)	30 (62.5)	22 (66.7)	40 (72.7)	27 (81.1)
Disease duration in years, mean (SD)	1.3 (0.9)		7.4 (5.2)	
Diffuse skin involvement, n (%)	48 (100)		37 (67.3)	
mRSS, mean (SD)	21.3 (8.7)		15.3 (10.4)	
Local skin score, mean (SD)	1.7 (0.8)		1.1 (0.9)	
FVC % predicted, mean (SD)	76.0 (19.8)		77.3 (19.8)	
SSc-associated autoantibody, n (%)*				
Antitopoisomerase I	12/41 (29.3)		15 (27.3)	
Anti-RNA polymerase III	17/38 (44.7)		17 (30.9)	
Anticentromere	1/36 (2.8)		7 (12.7)	
Mycophenolate, n (%)†	19 (39.6)		4 (7.3)	
Methotrexate, n (%)†	9 (18.8)		8 (14.5)	
Cyclophosphamide, n (%)†	1 (2.1)		1 (1.8)	

*Indicates positive among those recorded.
†Indicates patients taking medication at the time of biopsy.
FVC, forced vital capacity; GENISOS, Genetics versus Environment in Scleroderma Outcomes Study; HC, healthy control; mRSS, modified Rodnan Skin Score;PRESS, Prospective Registry for Early Systemic Sclerosis; SSc, systemic sclerosis.

was performed using the available GENISOS (n=55) and matched HC RNA samples (n=33) from that study. Differences in disease characteristics of the patients whose skin gene expression was analysed in GENISOS and PRESS are shown in table 1. On average, PRESS patients had a shorter disease duration at the time of biopsy than GENISOS patients (1.3 vs 7.4 years, respectively). Compared with GENISOS patients, PRESS patients had higher CD8 T cell, CD4 T cell, B cell and natural killer (NK) cell signatures in addition to M1 and M2 macrophage signatures (figure 2B). Fibroblast signatures were similar between the two

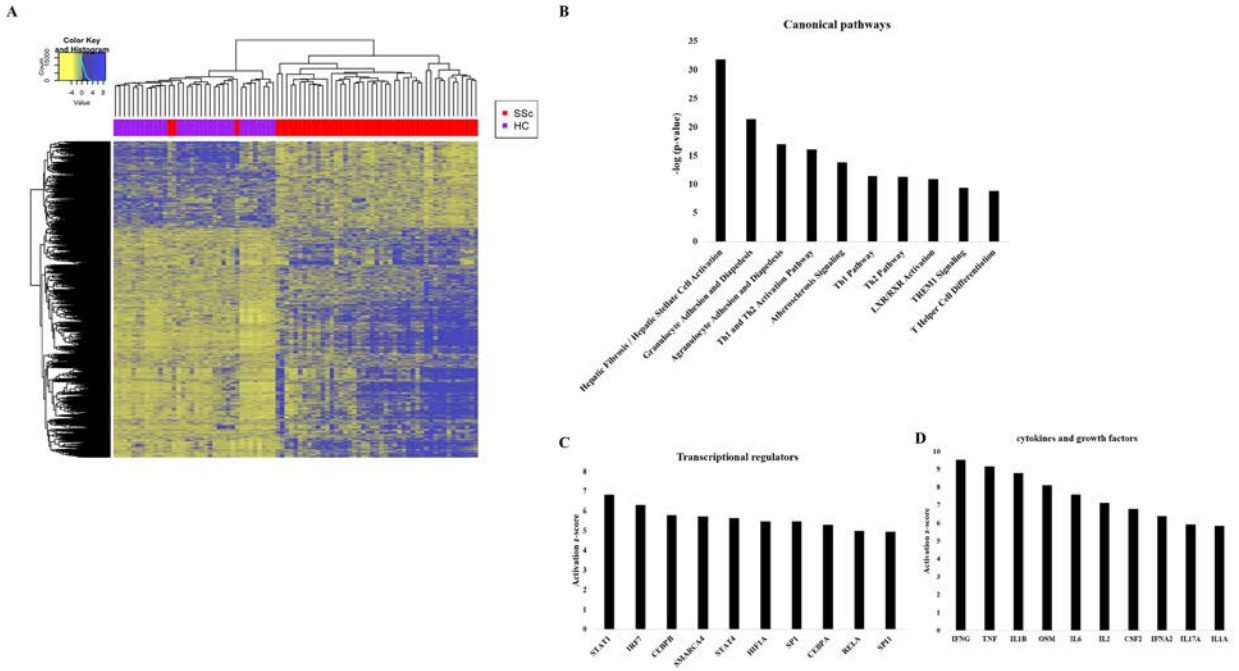


Figure 1 Differentially expressed transcripts and pathways in Prospective Registry for Early Systemic Sclerosis systemic sclerosis (SSc) patients compared with healthy controls (HCs). (A) Heatmap of differentially expressed transcripts, represented by z-score normalised count values. Unsupervised hierarchical clustering is shown at the top, with HCs represented by purple squares and SSc patients represented by red squares. (B) Top 10 over-represented pathways in SSc compared with HC as determined by Ingenuity Pathway Analysis of differentially expressed transcripts (fold change >1.5 or <0.67 in SSc vs HC, with false discovery rate <0.05). (C) Top 10 predicted upstream transcriptional regulators in SSc compared with HC. (D) Top 10 predicted upstream cytokines/growth factors.

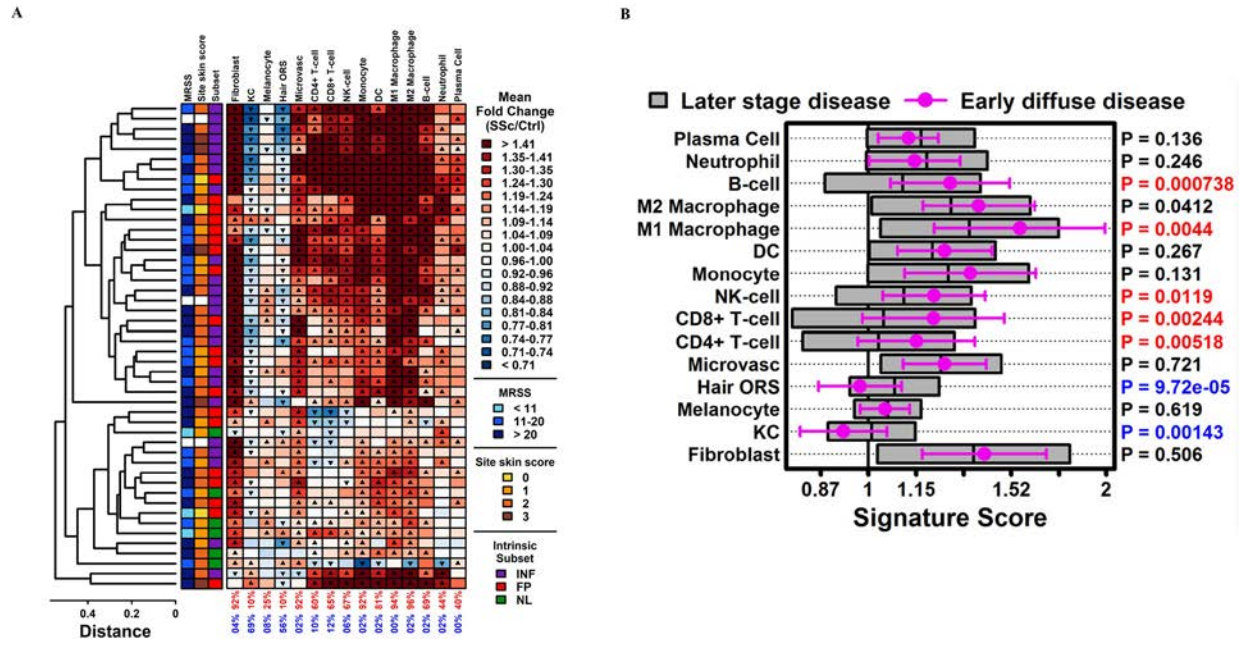


Figure 2 Cell type signatures in skin of PRESS SSc patients compared with healthy controls and compared with GENISOS SSc patients. (A) Cell type signature scores for each SSc sample (n=48). Scores represent the average fold-change (SSc/HC) for 125 cell type-specific signature genes (see online supplementary methods). Up-triangles indicate significantly higher scores for signature genes compared with non-signature genes ($p<0.05$, Wilcoxon rank-sum test). Down-triangles indicate significantly lower scores for signature genes compared with non-signature genes ($p<0.05$, Wilcoxon rank-sum test). Bottom margin values indicate the percentage of up-triangles (red) and down-triangles (blue), respectively. Patients were clustered based on signature scores (average linkage, Euclidean distance). The coloured boxes to the left of the cell type signature scores indicate the mRSS (left), local skin score at the site of the biopsy (middle) and the intrinsic subset classification, with legends at the right of the figure. White boxes (n=3) indicate no skin scores recorded at the time of the biopsy. (B) Signature scores for PRESS patients (n=48) were compared with those of GENISOS patients (n=55). The mean PRESS score is represented by round symbols with error bars spanning ±1 SD. The mean GENISOS score is represented by the midline for each grey box with boxes spanning ±1 SD. Right margin p values were obtained from a two-sample t-test of PRESS versus GENISOS scores (red: PRESS>GENISOS, FDR<0.05; blue: PRESS<GENISOS, FDR<0.05). Ctrl (HC), healthy control; DC, dendritic cell; GENISOS, Genetics versus Environment in Scleroderma Outcomes Study; hair ORS, hair outer root sheet; KC, keratinocyte; mRSS, modified Rodnan Skin Score; PRESS, Prospective Registry for Early Systemic Sclerosis; SSc, systemic sclerosis. FP, fibroproliferative subset; INF, inflammatory subset; NL, normal-like subset.

cohorts, while hair outer root sheet and keratinocyte signatures were lower in PRESS compared with GENISOS. Restricting the analysis to GENISOS patients with diffuse SSc and >3 years disease duration (n=28), PRESS patients had higher immune cell signatures, although the differences were smaller in this subgroup analysis (online supplementary figure 1). The prevalence of upregulated CD8 T cell, CD4 T cell and B cell signatures was relatively low in GENISOS as a whole (22%, 20% and 22%, respectively), including among the 28 patients with diffuse cutaneous involvement and >3 years disease duration (21%, 18% and 21%, respectively) (online supplementary figure 2).

To characterise clinical correlates of immune cell signatures within both cohorts, we pooled the data and performed multivariable regression analyses where the associations of disease duration, extent of skin involvement (as determined by mRSS), FVC % predicted and immunosuppression (comparing those on no immunosuppression to those on methotrexate, mycophenolate or cyclophosphamide at the time of biopsy) with cell type signatures (dependent variable) were examined. Adaptive immune cell signatures were inversely associated with disease duration after adjustment for mRSS, FVC % predicted and immunosuppression. By contrast, M1 and M2 macrophages and fibroblasts associated with mRSS but did not significantly associate with disease duration after adjustment for other clinical variables (table 2). These associations were similar after additional adjustment for PRESS versus GENISOS cohorts, suggesting that the observations were not driven by batch effects (online supplementary table 3). Of note, the investigated cell type signatures were not associated with immunosuppressive treatment in the univariable analysis (data not shown) or multivariable analysis (table 2).

Examination of available follow-up samples in the PRESS cohort

The majority of follow-up biopsies showed declines in immune cell signatures compared with their original biopsies (online supplementary figure 3A, B and online supplementary table 4). Fibroblast signatures were more variable at follow-up, with a small decline on average. Keratinocyte signatures were increased in most follow-up biopsies. Most of the patients with follow-up biopsies had a decline in mRSS from baseline to follow-up, and mRSS change correlated with changes in immune cell and fibroblast signatures numerically.

Histological associations with gene expression profiles

Paraffin-embedded skin biopsy samples concurrently collected from a subgroup of PRESS SSc patients were evaluated histologically using standard H&E staining and IHC staining for markers of macrophages (CD68, CD163, AIF1), endothelial cells (CD31) and myofibroblasts (α-smooth muscle actin(SMA)), as well as markers of adaptive immune cells CD3, CD4, CD8, CD20 and CD56 (it should be noted that CD4 is also expressed in monocytes/macrophages, although at a much lower intensity than in CD4 T cells,[19] and that CD56 is expressed in a subset

Table 2 Multivariable regression analyses of key clinical variables with cell type-specific signatures in pooled PRESS and GENISOS datasets

	Coefficient	95% CI	P value
CD8 T cell*			
Disease duration	−0.026	−0.042 to −0.009	<0.01
mRSS	0.006	−0.002 to 0.014	0.12
FVC % pred	−0.001	−0.005 to 0.002	0.54
No immunosuppression	0.095	−0.058 to 0.249	0.22
CD4 T cell*			
Disease duration	−0.02	−0.034 to −0.006	<0.01
mRSS	0.004	−0.003 to 0.010	0.25
FVC % pred	−0.001	−0.004 to 0.002	0.49
No immunosuppression	0.06	−0.069 to 0.190	0.36
NK cell*			
Disease duration	−0.019	−0.031 to −0.007	<0.01
mRSS	0.004	−0.001 to 0.010	0.12
FVC % pred	−0.001	−0.004 to 0.001	0.39
No immunosuppression	0.086	−0.026 to 0.197	0.13
B cell*			
Disease duration	−0.023	−0.037 to −0.009	<0.01
mRSS	0.002	−0.005 to 0.009	0.56
FVC % pred	−0.001	−0.004 to 0.002	0.5
No immunosuppression	−0.014	−0.146 to 0.119	0.84
M1 macrophage*			
Disease duration	−0.013	−0.030 to 0.004	0.13
mRSS	0.013	0.005 to 0.021	<0.01
FVC % pred	−0.002	−0.005 to 0.002	0.36
No immunosuppression	0.04	−0.119 to 0.199	0.62
M2 macrophage*			
Disease duration	−0.001	−0.014 to 0.012	0.91
mRSS	0.014	0.007 to 0.020	<0.01
FVC % pred	−0.001	−0.003 to 0.002	0.61
No immunosuppression	0.005	−0.117 to 0.127	0.94
Fibroblast*			
Disease duration	0.001	−0.015 to 0.016	0.93
mRSS	0.016	0.008 to 0.023	<0.01
FVC % pred	0.001	−0.002 to 0.004	0.57
No immunosuppression	0.028	−0.119 to 0.174	0.71

*Cell type transcript signature used as the dependent variable in the multivariable model.
FVC, forced vital capacity; GENISOS, Genetics versus Environment in Scleroderma Outcomes Study; mRSS, modified Rodnan Skin Score; NK, natural killer; PRESS, Prospective Registry for Early Systemic Sclerosis.

but not all NK cells). Demographics for these samples are shown in online supplementary table 5, and representative slides are shown in online supplementary figure 4. As expected, SSc skin had increased collagen thickness, α-SMA expression and macrophage markers compared with HC skin (online supplementary table 6). Markers of adaptive immune cells were also increased in SSc compared with HC skin (online supplementary table 7). Clinical correlates of IHC staining are shown in online supplementary table 8.

Importantly, cell type signature scores for macrophages and adaptive immune cells based on RNA sequencing data correlated with IHC staining for markers of macrophages and adaptive immune cells, respectively (table 3). Histologically, CD68 and CD163 tracked roughly in parallel, consistent with the reported difficulty in discerning M1 from M2 subtypes with these markers in human cells.[20] Taken together, the correlations with IHC staining support the validity of the gene expression-based cell type signatures. Moreover, fibroblast gene expression signature scores correlated with α-SMA (Spearman's rank order correlation coefficient 0.73, p<0.01) and collagen thickness (Spearman's rank order correlation coefficient 0.76, p<0.01).

Table 3 Correlation of immune cell gene expression signatures with immunohistochemical staining of immune cell markers

Cell abundance by IHC staining	Cell type signature score	Spearman's r (p value)
CD68	M1 macrophage	0.45 (0.02)
CD68	M2 macrophage	0.50 (0.01)
CD163	M1 macrophage	0.47 (0.02)
CD163	M2 macrophage	0.57 (<0.01)
AIF1	M1 macrophage	0.66 (<0.01)
AIF1	M2 macrophage	0.69 (<0.01)
CD3	CD4 T cell	0.61 (<0.01)
CD3	CD8 T cell	0.63 (<0.01)
CD4	CD4 T cell	0.49 (<0.01)
CD8	CD8 T cell	0.67 (<0.01)
CD20	B cell	0.54 (<0.01)
CD56	NK cell	0.24 (0.22)

IHC, immunohistochemical; NK, natural killer.

Association of cell type signature with disease course

A summary of mRSS, FVC and immunosuppression use 12 months after initial skin biopsy is shown in online supplementary table 9. 78.6% of patients were taking immunosuppressive medication 12 months after initial biopsy, which is expected for a cohort of early diffuse SSc patients. Cell type signatures did not significantly predict change in mRSS 6 or 12 months after biopsy, or change in FVC 12 months after biopsy (online supplementary table 10). Similarly, transcripts recently described as predictive of mRSS progression[21] based on samples collected in a phase II study of tocilizumab did not significantly predict postbiopsy mRSS change in this cohort (online supplementary figure 5). Restricting the analysis to those treated with immunosuppressive medications during follow-up also did not show predictive significance for the immune cell signatures (data not shown).

We then looked for associations with the preceding STPR, which was found to be an independent predictor of mortality in patients with early diffuse SSc.[15] Significant correlations were seen between immune cell signatures and STPR preceding the biopsy (figure 3 and online supplementary table 10). Thus, immune cell signatures in this cohort were associated with STPR up to the time of biopsy, but did not predict subsequent progression.

Comparison to intrinsic subset analysis

The PRESS samples were also assigned to one of four intrinsic subsets (inflammatory, fibroproliferative, limited or normal-like) using previously described methodology by Dr Whitfield's group.[17 18] Thirty-two out of 33 HCs were classified as normal-like, with 1 out of 33 classified as limited (data not shown). Among SSc patients, 23 were classified as inflammatory, 19 as fibroproliferative and 6 as normal-like (figure 2A and online supplementary figure 6). As shown in the figures, there was an over-representation of adaptive immunity cell type signatures in the inflammatory subset of patient samples.

Examination of longitudinal samples revealed that among five samples classified as inflammatory, follow-up biopsies from three of these individuals were classified in non-inflammatory subsets (two fibroproliferative and one normal-like), whereas

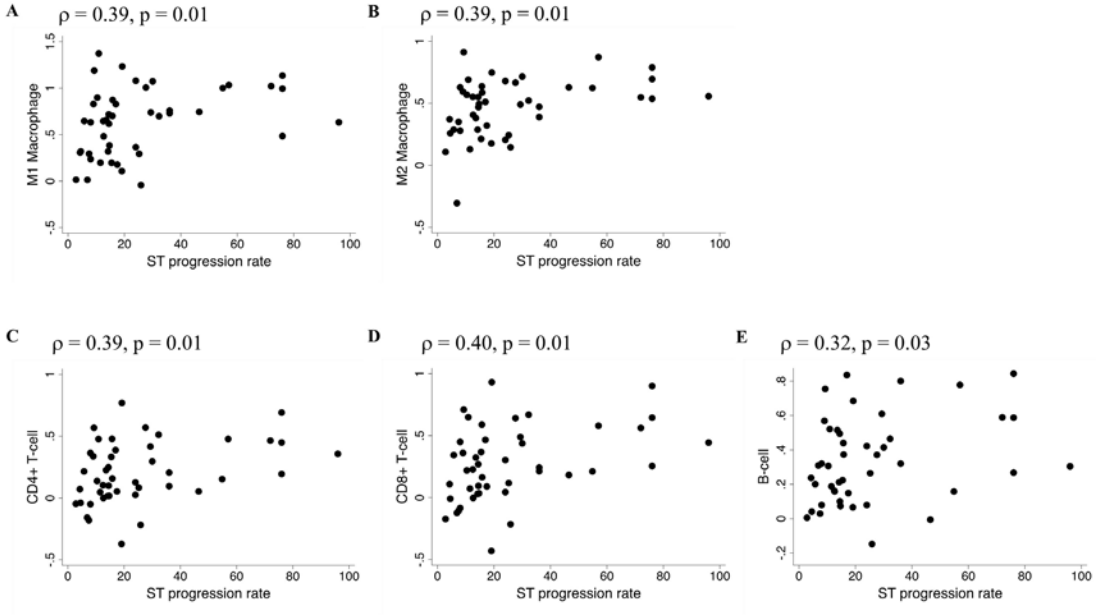


Figure 3 Associations between preceding skin thickness progression rate and skin immune cell type signatures in PRESS SSc patients. Skin thickness progression rate (mRSS at the time of biopsy/years since first skin thickening or puffy fingers) preceding the skin biopsy is plotted on the x-axis. Cell type signature scores for (A) M1 macrophages, (B) M2 macrophages, (C) CD4 cells, (D) CD8 T cells or (E) B cells are plotted on the y-axis. mRSS, modified Rodnan Skin Score; PRESS, Prospective Registry for Early Systemic Sclerosis; SSc, systemic sclerosis; ST, skin thickness.

none of the individuals with biopsies in the fibroproliferative or normal-like subsets on the initial biopsy had a follow-up biopsy in the inflammatory subset ().

Regarding mRSS course, the intrinsic subsets did not significantly predict mRSS change 6 or 12 months postbiopsy in the overall cohort or in the subgroup of patients taking immunosuppressive agents during follow-up (online supplementary table 11). In agreement with the immune cell signature data, STPR preceding the biopsy was significantly higher in the inflammatory subset (online supplementary table 12).

DISCUSSION

Histological and gene expression analyses have demonstrated variable degrees of innate and adaptive immune cells in affected SSc skin.[5–11 22–28] In the current study, we measured whole transcriptome expression and cell type signatures in skin specimens in a large cohort of patients specifically with early diffuse SSc and matched HCs. More than half of patients in this cohort had upregulation of CD8 T cell, CD4 T cell and B cell signatures, a higher prevalence than what was observed in patients with longer disease duration from the GENISOS cohort. We also observed a higher prevalence of M1 and M2 macrophage signatures in the skin of early diffuse SSc patients. In patients with longitudinally collected biopsies, immune cell signatures declined on average from initial to follow-up biopsies. These results parallel the clinical observation that early SSc has an edematous, inflammatory phase followed by a more fibrotic phase, and the histological findings in SSc showing an early 'cellular stage' characterised by cellular infiltrates in the dermis followed by a later 'fibrotic stage' characterised by increased collagen deposition.[22 23]

Multivariable regression analysis including all samples from the PRESS and GENISOS cohorts showed that adaptive immune cell signatures were significantly associated with shorter disease duration even after adjustment for immunosuppression, severity of skin disease (as assessed by mRSS) and lung disease (as assessed by FVC), whereas macrophage and fibroblast signatures associated predominately with mRSS. These results suggest that the determinants of adaptive versus innate immune cell infiltration in the skin may differ. This can also have implications for target population enrichment strategies in clinical trials, although the observation needs to be confirmed in future studies.

Ingenuity Pathway Analysis suggested that inflammatory cytokines had a more prominent role in driving the dysregulated gene expression in early diffuse SSc compared with later stage disease. Of note, the vast majority of early diffuse SSc patients with a fibroblast signature had a concomitant M1 and/or M2 macrophage signature, and many had concomitant adaptive immune cell signatures, suggesting co-occurrence of dysregulated fibroblast and immune cell function in a majority of early diffuse SSc patients. Our gene expression and IHC data add to the large body of evidence that macrophages are upregulated in SSc.[29 30] Macrophages are capable of detecting innate immune stimuli and producing both pro-inflammatory and pro-fibrotic cytokines, including some (eg, IL-6 and TGFβ) that are implicated in SSc pathogenesis. However, the effects of macrophages within the skin and other end organs in SSc require further study.

Taken together, our results indicate that innate and adaptive immune cell activity in the skin is a prominent feature of early diffuse SSc. TGFβ, a key pro-fibrotic cytokine implicated in SSc pathogenesis,[31] appears to have a less prominent role in driving the dysregulated gene expression observed during this early, inflammatory phase, in contrast to its prominent role in later-stage disease.

Histological scoring in concurrently collected skin samples supported the gene expression data, demonstrating upregulation of macrophage, adaptive immune cell and fibrotic markers. Immune cell markers correlated with their respective gene expression signatures, and fibrosis markers correlated with fibroblast gene expression signatures. These results support the validity of the gene expression-based cell type signatures.

The RNA processing method used here (ribosomal RNA reduction) enabled the provision of an unbiased comprehensive list of differentially expressed lncRNAs, because this method (unlike poly (A) enrichment) does not remove lncRNAs that do

not have a poly(A) tail.[32] We have provided a list of differentially expressed lncRNAs expressed in the skin of early diffuse SSc compared with HC, as well as their associations with mRSS. Although our currently available pathway and predicted upstream regulator analytic methods do not include analysis of lncRNAs, the list of disease-relevant lncRNAs represents a resource for follow-up mechanistic studies in this novel area of research.

The carefully collected clinical data in the well-phenotyped PRESS cohort enabled us to examine the correlation of the SSc gene expression profile with the progression rate of skin fibrosis prior to and following skin biopsy. Immune cell signatures were associated with preceding STPR, while they did not have predictive significance for postbiopsy mRSS change. Similarly, transcripts found to be predictive of mRSS progression in previous work[21] were not significantly associated with postbiopsy mRSS change in this study. Intrinsic subset classification (normal-like, inflammatory and fibroproliferative)[18] did not show predictive significance for mRSS change 6 or 12 months after biopsy. These findings suggest that the use of these previously described gene signatures and subsets for predicting changes in mRSS may not be generalisable to all cohorts. Further research will be needed to determine whether or not a model for prediction of disease progression based on skin gene expression can be universally applied across cohorts, particularly in patients on treatment with commonly used immunosuppressive medications typified in PRESS. The data in this study suggest that skin gene expression signatures in early diffuse SSc are more of a reflection of preceding skin thickness progression than predictors of subsequent progression, supporting the notion that skin is a prominent end organ in SSc rather than an effector organ that drives disease progression.

Our study has several strengths. We examined the transcript expression profile of a relatively large number of skin samples in a well-phenotyped early diffuse SSc cohort using a sensitive, comprehensive RNA sequencing method and compared the results to a later stage SSc transcript expression dataset generated using the same technology. The gene expression-based cell type signatures were validated by IHC staining in concurrently collected samples. There were some limitations to this study that merit discussion. Only a small subgroup of patients (n=8) had follow-up samples available, limiting the ability to analyse changes in gene expression during disease progression. Our future studies will focus on longitudinal collection of early diffuse SSc skin samples. As is common in observational studies and most previous SSc skin gene expression studies, patients enrolled in PRESS were treated according to the standard of care, with the majority being treated with mycophenolate mofetil or methotrexate, which might have affected skin transcript expression.

In conclusion, this large-scale analysis of whole transcriptome expression in the skin of early diffuse SSc patients revealed a high prevalence of both innate and adaptive immune cell activity. Immune cell signatures were associated with preceding STPR but were not predictive of subsequent mRSS progression. These results shed light on the early pathogenesis of diffuse SSc and could have implications for clinical trials targeting the immune system in SSc patients.

Author affiliations
[1]Division of Rheumatology and Clinical Immunogenetics, University of Texas Health Science Center at Houston, Houston, Texas, USA
[2]Scleroderma Program, Department of Internal Medicine, Division of Rheumatology, University of Michigan, Ann Arbor, Michigan, USA
[3]Ohio University Heritage College of Osteopathic Medicine, Athens, Ohio, USA
[4]Department of Internal Medicine, The Jewish Hospital, Cincinnati, Ohio, USA
[5]Department of Medicine, Section of Allergy, Rheumatology, and Immunology, Yale University, New Haven, Connecticut, USA
[6]Division of Rheumatology, Department of Internal Medicine, University of Utah and Salt Lake Regional Veterans Affairs Medical Center, Salt Lake City, Utah, USA
[7]Division of Rheumatology, Department of Medicine, MedStar Georgetown University Hospital, Washington, DC, USA
[8]Division of Rheumatology and Immunology, Department of Medicine, Medical University of South Carolina, Charleston, South Carolina, USA
[9]Department of Rheumatology, Hospital for Special Surgery, New York City, New York, USA
[10]Division of Rheumatology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA
[11]School of Biomedical Informatics, University of Texas Health Science Center at Houston, Houston, Texas, USA
[12]Department of Microbiology, Section of Rheumatology, Boston University School of Medicine, Boston, Massachusetts, USA
[13]Boehringer Ingelheim Pharmaceuticals Inc, Ridgefield, Connecticut, USA
[14]Boehringer Ingelheim International GmbH, Biberach, Germany
[15]Department of Biomedical Data Science, Dartmouth College Geisel School of Medicine, Lebanon, New Hampshire, USA
[16]Department of Molecular and Systems Biology, Dartmouth College Geisel School of Medicine, Hanover, New Hampshire, USA
[17]Division of Rheumatology, George Washington University School of Medicine and Health Sciences, Washington, DC, USA
[18]Division of Rheumatology and Clinical Immunology, Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA
[19]Division of Rheumatology, Massachusetts General Hospital, Boston, Massachusetts, USA
[20]Division of Rheumatology, Vagelos College of Physicians and Surgeons, New York City, New York, USA

Twitter Shervin Assassi @ShervinAssassi

Acknowledgements We acknowledge Yoko Takata and Jianjun Shen (MD Anderson Cancer Center, Houston, TX) for performing RNA sequencing. JS is supported by a CPRIT Core Facility Support Award RP170002. We acknowledge Melissa Stephens (UT Health Science Center, Houston) for performing some of the immunohistochemical staining. We thank Sam Theodore (UT Health Science Center, Houston) for assistance with skin biopsies and database management. Part of this work was previously presented orally at the American College of Rheumatology meeting in 2017 (Arthritis Rheumatol. 2017;69 (suppl 10)).

Contributors SA, BS, DK, SV and TF designed the study. BS, DK, MEH, TF, VS, FH, JG, AAS, JB, AMSB, MW, SV, PB, NW, JC and SA were involved in acquisition of data. All authors were involved in analysis and interpretation of the data. All authors were involved in manuscript preparation and approved the final version of the manuscript.

Funding BS is supported by a grant from the Arthritis National Research Foundation. DK is supported by a Scleroderma Foundation SCORE grant and NIH/NIAMS K24 AR 063120. The PRESS Data Coordinating Center at University of Michigan was supported by NIH/NIAMS K24 AR063120 to DK. SA is supported by a Scleroderma Foundation SCORE grant, a DoD grant W81XWH-61-1-0296 and by NIH/NIAMS R01AR073284. MW is supported by a grant from the Scleroderma Foundation. This work was partly supported by the UT Houston Bioinformatics and High Performance Computing Service Center and CPRIT grant RP170668 (WJZ). WJZ is supported by NIH 1UL1 TR003167-01. JMF and MLW are supported by the Scleroderma Research Foundation, Dr Ralph and Marian Falk Medical Research Trust Catalyst and Transformational Award and NIH/NIAMS P50 AR060780-01. JMF is supported by NIH 5T32LM012204-03. EJB is supported by NIH/NIAMS K23 AR075112.

Competing interests DK reports consultancy fees from Acceleron, Actelion, Bayer, Blade Therapeutics, BMS, Galapagos, Genentech/Roche, GSK, Mitsubishi Tanabi, Sanofi-A ventis/Genzyme, and UCB Pharma, and reports grants from Bayer, BMS, and Genentech/Roche outside the submitted work, and reports ownership interest in Eicos Sciences, Inc. WRS reports consulting fees from UT Health Science Center at Houston. JKG reports grants from Corbus Pharmaceuticals, Cumberland Pharmaceuticals and Eicos Pharmaceuticals outside the submitted work. AAS reports support from Sanofi outside the submitted work. MLW reports grants and personal fees from Celdara Medical, LLC and personal fees from Abbvie, Acceleron, BMS, Corbus Pharmaceuticals and Boehringer Ingelheim outside the submitted work. JLB reports grants from Hoffman La Roche and the Scleroderma Foundation outside the submitted work. SB and PB are employees of Boehringer Ingelheim. FVC reports personal fees from Boehringer Ingelheim outside the submitted work. EJB reports a grant from Pfizer and support from Bohringer Ingelheim, Corbus Pharmaceuticals and Eicos outside the submitted work. MDM reports personal fees from Medtelligence, Actelion Pharma, Astellas, Mitsubishi-Tanabe, and Boehringer Ingelheim and grants from Bayer, Reata, Sanofi, Corbus, Eicos and Boehringer Ingelheim outside the submitted work. SA reports personal fees from Boehringer

Ingelheim and grants from Boehringer Ingelheim, Bayer and Momenta outside the submitted work.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement RNA sequencing data are available at NCBI's Gene Expression Omnibus, accession GSE130955.

ORCID iDs
Brian Skaug http://orcid.org/0000-0002-6689-4655
Dinesh Khanna http://orcid.org/0000-0003-1412-4453
Minghua Wu http://orcid.org/0000-0002-0388-2032
Sudha Visvanathan http://orcid.org/0000-0002-1170-9272
Jennifer M Franks http://orcid.org/0000-0003-2400-5431
Shervin Assassi http://orcid.org/0000-0002-8059-9978

REFERENCES

1 Denton CP, Khanna D. Systemic sclerosis. *Lancet* 2017;390:1685–99.
2 Varga J, Trojanowska M, Kuwana M. Pathogenesis of systemic sclerosis: recent insights of molecular and cellular mechanisms and therapeutic opportunities. *J Scleroderma Relat Disord* 2017;2:137–52.
3 Assassi S, Mayes MD. What does global gene expression profiling tell us about the pathogenesis of systemic sclerosis? *Curr Opin Rheumatol* 2013;25:686–91.
4 Martyanov V, Whitfield ML. Molecular stratification and precision medicine in systemic sclerosis from genomic and proteomic data. *Curr Opin Rheumatol* 2016;28:83–8.
5 Whitfield ML, Finlay DR, Murray JI, *et al*. Systemic and cell type-specific gene expression patterns in scleroderma skin. *Proc Natl Acad Sci U S A* 2003;100:12319–24.
6 Gardner H, Shearstone JR, Bandaru R, *et al*. Gene profiling of scleroderma skin reveals robust signatures of disease that are imperfectly reflected in the transcript profiles of explanted fibroblasts. *Arthritis Rheum* 2006;54:1961–73.
7 Milano A, Pendergrass SA, Sargent JL, *et al*. Molecular subsets in the gene expression signatures of scleroderma skin. *PLoS One* 2008;3:e2696.
8 Pendergrass SA, Lemaire R, Francis IP, *et al*. Intrinsic gene expression subsets of diffuse cutaneous systemic sclerosis are stable in serial skin biopsies. *J Invest Dermatol* 2012;132:1363–73.
9 Hinchcliff M, Huang C-C, Wood TA, *et al*. Molecular signatures in skin associated with clinical improvement during mycophenolate treatment in systemic sclerosis. *J Invest Dermatol* 2013;133:1979–89.
10 Assassi S, Swindell WR, Wu M, *et al*. Dissecting the heterogeneity of skin gene expression patterns in systemic sclerosis. *Arthritis Rheumatol* 2015;67:3016–26.
11 Lofgren S, Hinchcliff M, Carns M, *et al*. Integrated, multicohort analysis of systemic sclerosis identifies robust transcriptional signature of disease severity. *JCI Insight* 2016;1:e89073.
12 Moon S-J, Bae JM, Park K-S, *et al*. Compendium of skin molecular signatures identifies key pathological features associated with fibrosis in systemic sclerosis. *Ann Rheum Dis* 2019;78:817–25.
13 Frech TM, Shanmugam VK, Shah AA, *et al*. Treatment of early diffuse systemic sclerosis skin disease. *Clin Exp Rheumatol* 2013;31:166–71.
14 van den Hoogen F, Khanna D, Fransen J, *et al*. 2013 classification criteria for systemic sclerosis: an American College of rheumatology/European League against rheumatism collaborative initiative. *Ann Rheum Dis* 2013;72:1747–55.
15 Domsic RT, Rodriguez-Reyna T, Lucas M, *et al*. Skin thickness progression rate: a predictor of mortality and early internal organ involvement in diffuse scleroderma. *Ann Rheum Dis* 2011;70:104–9.
16 Swindell WR, Johnston A, Voorhees JJ, *et al*. Dissecting the psoriasis transcriptome: inflammatory- and cytokine-driven gene expression in lesions from 163 patients. *BMC Genomics* 2013;14:527.
17 Franks JM, Cai G, Whitfield ML. Feature specific quantile normalization enables cross-platform classification of molecular subtypes using gene expression data. *Bioinformatics* 2018;34:1868–74.
18 Franks JM, Martyanov V, Cai G, *et al*. A machine learning classifier for assigning individual patients with systemic sclerosis to intrinsic molecular subsets. *Arthritis Rheumatol* 2019;71:1701–10.
19 Wood GS, Warner NL, Warnke RA. Anti-Leu-3/T4 antibodies react with cells of monocyte/macrophage and Langerhans lineage. *J Immunol* 1983;131:212–6.
20 Barros MHM, Hauck F, Dreyer JH, *et al*. Macrophage polarisation: an immunohistochemical approach for identifying M1 and M2 macrophages. *PLoS One* 2013;8:e80908.
21 Stifano G, Sornasse T, Rice LM, *et al*. Skin gene expression is prognostic for the trajectory of skin disease in patients with diffuse cutaneous systemic sclerosis. *Arthritis Rheumatol* 2018;70:912–9.
22 Fleischmajer R, Perlish JS, Reeves JR. Cellular infiltrates in scleroderma skin. *Arthritis Rheum* 1977;20:975–84.
23 Fleischmajer R, Gay S, Meigel WN, *et al*. Collagen in the cellular and fibrotic stages of scleroderma. *Arthritis Rheum* 1978;21:418–28.
24 Roumm AD, Whiteside TL, Medsger TA, *et al*. Lymphocytes in the skin of patients with progressive systemic sclerosis. quantification, subtyping, and clinical correlations. *Arthritis Rheum* 1984;27:645–53.
25 Ishikawa O, Ishikawa H. Macrophage infiltration in the skin of patients with systemic sclerosis. *J Rheumatol* 1992;19:1202–6.
26 Kräling BM, Maul GG, Jimenez SA. Mononuclear cellular infiltrates in clinically involved skin from patients with systemic sclerosis of recent onset predominantly consist of monocytes/macrophages. *Pathobiology* 1995;63:48–56.
27 Hinchcliff M, Toledo DM, Taroni JN, *et al*. Mycophenolate mofetil treatment of systemic sclerosis reduces myeloid cell numbers and attenuates the inflammatory gene signature in skin. *J Invest Dermatol* 2018;138:1301–10.
28 Bosello S, Angelucci C, Lama G, *et al*. Characterization of inflammatory cell infiltrate of scleroderma skin: B cells and skin score progression. *Arthritis Res Ther* 2018;20.
29 Brown M, O'Reilly S. The immunopathogenesis of fibrosis in systemic sclerosis. *Clin Exp Immunol* 2019;195:310–21.
30 Toledo DM, Pioli PA. Macrophages in systemic sclerosis: novel insights and therapeutic implications. *Curr Rheumatol Rep* 2019;21:31.
31 Lafyatis R. Transforming growth factor β–at the centre of systemic sclerosis. *Nat Rev Rheumatol* 2014;10:706–19.
32 Kukurba KR, Montgomery SB. RNA sequencing and analysis. *Cold Spring Harb Protoc* 2015;2015:951–69.

EPIDEMIOLOGICAL SCIENCE

First external validation of sensitivity and specificity of the European League Against Rheumatism (EULAR)/American College of Rheumatology (ACR) classification criteria for idiopathic inflammatory myopathies with a Japanese cohort

Masatoshi Jinnin ,[1] Akiko Ohta,[2] Shoichiro Ishihara,[3] Hirofumi Amano,[4] Tatsuya Atsumi,[5] Manabu Fujimoto,[6] Takashi Kanda,[7] Yasushi Kawaguchi,[8] Atsushi Kawakami,[9] Akio Mimori,[10] Tsuneyo Mimori,[11] Toshihide Mimura,[12] Yoshinao Muro ,[13] Hajime Sano,[14] Jun Shimizu,[15] Tsutomu Takeuchi,[16] Yoshiya Tanaka ,[17] Kazuhiko Yamamoto ,[18] Takayuki Sumida,[19] Hitoshi Kohsaka,[20] The Research Team for Autoimmune Diseases, The Research Program for Intractable Disease of the Ministry of Health, Labor and Welfare

Handling editor Josef S Smolen

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumds-2019-215488).

For numbered affiliations see end of article.

Correspondence to
Dr Hitoshi Kohsaka, Tokyo, Japan;
kohsaka-tmd@umin.ac.jp

Received 7 April 2019
Revised 18 October 2019
Accepted 18 October 2019
Published Online First 6 November 2019

To cite: Jinnin M, Ohta A, Ishihara S, *et al*. *Ann Rheum Dis* 2020;**79**:387–392.

ABSTRACT

Objective To externally validate the performance of the new European League Against Rheumatism (EULAR)/American College of Rheumatology (ACR) classification criteria set for idiopathic inflammatory myopathies (IIM) with a Japanese cohort.

Methods This study included 420 IIM and 402 non-IIM cases. Probability of having IIM in each patient was calculated using the collected data set. The cut-off probability was set at 55%, as recommended by EULAR/ACR. Patients classified as IIM by the criteria were further subclassified with classification trees.

Results When the probability cut-off was set at 55%, the sensitivity/specificity of the new criteria to diagnose IIM were 89.3%/91.0% in the total cohort, 88.1%/95.1% without muscle biopsy data and 90.4%/65.5% with biopsy data. The cohort included 12 overlap syndrome patients with biopsy data, who were included as non-IIM cases in accordance with traditional Japanese methods. When they were included in the IIM cases, the specificity in patients with biopsy increased to 74.4%. The sensitivity/specificity of the new criteria to diagnose polymyositis/dermatomyositis (PM/DM) plus juvenile and amyopathic DM in the Japanese cohort was 87.4%/92.4%, which were greater than those of the Tanimoto's criteria revised to enable classification of amyopathic DM (ADM) (71.2%/87.8%) and were comparable with those of Bohan & Peter's criteria to diagnose those diseases except for ADM (88.4%/88.3%).

Conclusions Our study externally validated high specificity of the new criteria for the first time, although with several limitations, including low percentage of child patients. The new criteria have higher sensitivity and/or specificity in classification of PM/DM than the previously reported criteria, demonstrating its usefulness for interethnic patients.

Key messages

What is already known about this subject?

► New classification criteria for idiopathic inflammatory myopathies were proposed on the basis of the data analyses by the International Myositis Classification Criteria Project. Examined for sensitivity validation with external cohorts, they were recently approved by European League Against Rheumatism and American College of Rheumatology.

What does this study add?

► Our study externally validated high specificity of the new criteria for the first time.

How might this impact on clinical practice or future developments?

► The new criteria have higher sensitivity and/or specificity in classification of polymyositis/dermatomyositis than the previously reported criteria, demonstrating its usefulness for interethnic patients. Setting probability cut-off at 55% in the new criteria together with use of the classification tree was acceptable for the Japanese cohort.

INTRODUCTION

Idiopathic inflammatory myopathies (IIM) are heterogeneous disorders characterised by muscle weakness and muscle inflammation and include polymyositis (PM), dermatomyositis (DM) and inclusion body myositis (IBM).[1] Immune-mediated necrotising myopathy (IMNM) could be an independent disease entity in IIM. Various classification criteria for IIM and its subgroups have been published since 1970.[2] Among them, the Bohan & Peter's criteria and Tanimoto's criteria for PM/DM have been commonly used in Japan.[3–5]

Bohan & Peter's criteria, described in 1975, include five variables. There are some limitations, however; IBM and muscular dystrophies with

Table 1 Demographic characteristics of IIM cases and non-IIM cases in the study population

	IIM		Non-IIM	
	n=420		n=402	
Sex, n (%)				
Male	127	(30.2)	106	(27.9)
Female	293	(69.8)	296	(72.1)
Sex ratio (Female/Male)	2.31		2.79	
Age at onset, n (%)				
<18	6	(1.4)	17	(4.2)
18–40	71	(16.9)	144	(35.8)
≤40	341	(81.2)	236	(58.7)
NA*	2	(0.5)	5	(1.2)
Age at onset, mean (SD)	53.9	(15.4)	46.4	(19.2)
Age at diagnosis, n (%)				
<18	5	(1.2)	12	(3.0)
18–40	69	(16.4)	130	(32.3)
≤40	346	(82.4)	260	(64.7)
Age at diagnosis, mean (SD)	55.1	(15.4)	48.8	(18.9)
Calendar year of diagnosis, n (%)				
2007	8	(1.9)	12	(3.0)
2008	17	(4.0)	28	(7.0)
2009	45	(10.7)	41	(10.2)
2010	80	(19.0)	52	(12.9)
2011	111	(26.4)	91	(22.6)
2012	142	(33.8)	157	(39.1)
2013	17	(4.0)	21	(5.2)

Clinical data other than diagnostic information were shown: details of diagnostic information are shown in online supplementary table S4.
*NA: information not available.
IIM, idiopathic inflammatory myopathies.

inflammatory changes cannot be excluded. In addition, the skin rashes are not specified.

Tanimoto's criteria, published in 1995, include nine variables. Confusingly, definitions of the Gottron's sign and linear extensor erythema are different from those in other criteria. Only anti-Jo-1 antibody is included as the myositis-specific autoantibody, not more recently identified antibodies. Furthermore, because Tanimoto's criteria rely mainly on muscle abnormalities and related systemic rheumatic symptoms, amyopathic DM (ADM) cannot be diagnosed. Since the criteria have been used to identify patients for the purpose of subsidisation of medical expenses in Japan, there were some problems, especially among patients with ADM with severe interstitial lung diseases, until the recent revision in 2015 (online supplementary table 1).

Unlike previously published criteria, which were defined primarily on the basis of clinical impressions of experts and without validation by external cohorts, a new set of criteria were proposed on the basis of data analyses by the International Myositis Classification Criteria Project (IMCCP). Demographic, clinical and laboratory data of the 976 patients with IIM and consisting of 75 variables were collected and compared with those of 624 comparators which included patients with a broad spectrum of mimicking conditions, such as autoimmune myopathy or non-inflammatory myopathy. Data analysis revealed that 16 variables could distinguish the IIM cases from the comparators most effectively. The new criteria were examined for sensitivity validation with external cohorts and recently approved by European League Against Rheumatism (EULAR) and American College of Rheumatology (ACR) (http://www.imm.ki.se/biostatistics/calculators/iim).[6 7]

The new criteria have three distinct features, and the first is that they were developed differentially for patients who underwent muscle biopsy and those who did not. This consideration is primarily for juvenile cases and made the criteria applicable to adult and juvenile patients. The second feature of the new criteria is that each variable of the criteria was assigned a score based on its ability to discriminate IIM from the comparators. The criteria provide total aggregated scores for each patient, and the probability of IIM in each patient can be calculated from the total aggregated score using a formula. Last, while the original paper suggested a recommended cut-off probability for IIM as 55%,[6 7] there is users' flexibility in the setting of cut-offs according to the type of study; for example, the cut-off points can be increased to facilitate high specificity for clinical trials that should require stringent classification.

Taken together, IMCCP concluded that at present the new EULAR/ACR-approved criteria are the most reliable. Yet to be done for the criteria are validation of specificity with external cohorts and also for cases with ethnicities that may be underrepresented in the original cohort. Accordingly, in the current study, we validated the specificity as well as the sensitivity of the new criteria using a Japanese cohort. This is one of the research projects of subcommittee for PM/DM in the Japanese National Research Committee on Autoimmune Diseases supported by the Ministry of Health, Labour and Welfare of Japan. We discuss the usefulness and limitations of the criteria when used for Japanese patients with IIM.

MATERIALS AND METHODS

Study subjects and data collection

The subjects of the present study include 420 IIM cases and 402 non-IIM cases with definitive diagnosis of IIM or other diseases made by experts including the authors at 19 institutes participating in this national project between 2007 and 2013. The project was approved by research ethics committees of individual institutes and carried out by expert rheumatologists, neurologists, dermatologists and epidemiologists in these institutes. As inclusion criteria, the clinical data of allocated number of IIM cases assigned to each facility were retrospectively collected in the order of new patient visits from medical records using questionnaires (online supplementary table 2), which were distributed to the experts. There were no exclusion criteria.

Non-IIM cases were also collected from each institute, in the order of patient visits (online supplementary table 3). Diagnosis of all cases was validated by an expert rheumatologist, neurologist and dermatologist.

Methodology of new EULAR/ACR classification

The new EULAR/ACR classification criteria show the probability of a patient having IIM for use in clinical and research settings.[6 7] In brief, the new criteria consist of the muscle findings, skin findings and laboratory measures.[6 7] Each item is assigned a weighted score, and the aggregated scores are calculated by summing the score points. The aggregated scores can be converted into a probability of IIM using the following formulas:

Probability of IIM = 1/[1 + exponential (6.49 – score)] when muscle biopsy data are present or

Probability of IIM = 1/[1 + exponential (5.33 – score)] when muscle biopsy data are absent

Unknown or missing data were regarded as 'score 0'. Among the information included in the questionnaires, only the criteria items were used for classification of patients and comparators. The web-based calculator and the Excel file that can be downloaded from criteria web page were not used in this study. The correlation between total aggregated scores and varying probability of having

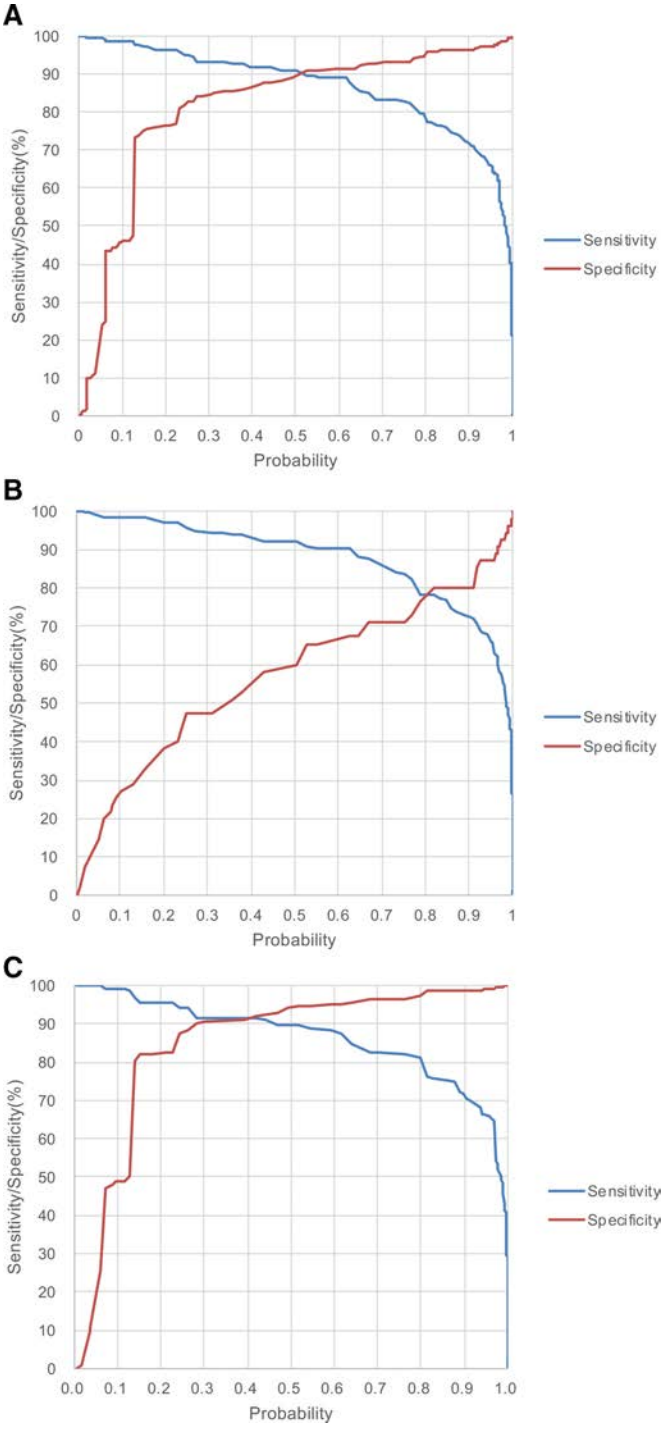


Figure 1 Sensitivity/specificity and probability of the EULAR/ACR classification criteria for IIM in the all Japanese cases (A), the cases without muscle biopsy data (B) and the cases with muscle biopsy data (C). Blue line and red lines indicate sensitivity and specificity, respectively. ACR, American College of Rheumatology; EULAR, European League Against Rheumatism; IIM, idiopathic inflammatory myopathies.

IIM are shown in the previous paper.[6 7] A patient is classified as an IIM case if the patient's probability is above a specified cut-off value.

Furthermore, a patient classified as IIM by the EULAR/ACR classification criteria can be further subclassified by a classification tree into six groups: PM, DM, ADM, juvenile DM (JDM), other juvenile myositis and IBM.[6 7]

Other methods are described in online supplementary text.

RESULTS

Demographic characteristics of IIM and non-IIM cases in the study population

Data from 420 IIM and 402 non-IIM cases were collected at the 19 institutes that participated in the present study. The IIM cases consisted of patients with five subgroups: PM (n=137; three patients had IMNM), DM (n=188), ADM (n=64), JDM (n=6) and IBM (n=7) (online supplementary table 4). The non-IIM cases consisted of patients with 31 different diseases including systemic lupus erythematosus (SLE; n=88), systemic vasculitis (n=63), systemic sclerosis (n=60), mitochondrial myopathy (n=8) and contact dermatitis (n=9).

The IIM cases consisted of 127 male and 293 female patients, whereas the non-IIM cases consisted of 106 male and 296 female patients (table 1). The mean age of onset was 53.9±15.4 years in the IIM cases and 46.4±19.2 years in the non-IIM cases, whereas the mean ages at diagnosis was 55.1±15.4 years in IIM and 48.8±18.9 years in non-IIM. The number of cases under 18 years of age at diagnosis was five (1.2%) in IIM and 12 (3.0%) in non-IIM.

The frequency of each variable in IIM cases and non-IIM cases are shown in online supplementary table 5. L2 (elevated serum levels of lactate dehydrogenase (LDH)) was the most common (93.8%) in patients with IIM among items included in the new criteria. Among the other items, skin biopsy was performed in 174 (41.4%) out of 420 IIM and 87 (21.6%) out of 402 non-IIM. Electromyogram and muscle MRI were performed in 67.9% and 70.2% of patients with IIM, respectively.

Sensitivity and specificity of the EULAR/ACR classification criteria for the Japanese cohort

The distribution of the probabilities of both IIM and non-IIM in all Japanese cases, the cases without muscle biopsy data and the cases with muscle biopsy data are shown in online supplementary figure 1.

The IMCCP team concluded that the best balance between sensitivity and specificity to diagnose IIM was found at probability of 55% in patients both with and without muscle biopsy data.[6 7] As was done by the team, sensitivity and specificity of the new criteria in the all Japanese cohort for varying probability cut-offs were calculated (figure 1A). Those in the patients with IIM without the muscle biopsy data and with the muscle biopsy data were also calculated separately (figure 1B, C). The optimal probability cut-off for sensitivity and specificity was found at 50% in all Japanese cases; sensitivity and specificity were 91.0% (95% CI: 87.79 to 93.52) and 89.8% (86.42–92.58), respectively. The optimal cut-offs were found at 40% in patients without muscle biopsy data; sensitivity and specificity were 91.6% (86.87–95.02) and 91.9% (88.55–94.57), respectively and at 65% in patients with muscle biopsy data; sensitivity and specificity were 87.6% (82.49–91.68) and 70.9% (57.10–82.37), respectively.

When the cut-off of 55% (recommended by the IMCCP team) was applied, the sensitivity and specificity in the all Japanese patients were 89.3% and 91.0%, respectively (table 2). In patients without muscle biopsy data, they were 88.1% and 95.1% respectively, and in patients with muscle biopsy data they were 90.4% and 65.5%, respectively (table 2). Since shift of the probability cut-off from the optimal values to 55% makes subtle changes (<6%) in the sensitivity and specificity, 55% should also be acceptable for Japanese patients. The new criteria were therefore validated with a Japanese cohort. Receiver operating characteristics curve analysis indicated that the area under the curve for all Japanese cases, the cases without muscle biopsy data and the cases with muscle

Table 2 Sensitivity and specificity of EULAR/ACR classification criteria for IIM

	No. of subjects			No. of positive cases		Sensitivity	Specificity
	Total	IIM	Non-IIM	IIM	Non-IIM	% (95% CI)	% (95% CI)
EULAR/ACR classification criteria for IIM							
Total	822	420	402	375	36	89.3 (85.93 to 92.08)	91.0 (87.82 to 93.65)
Without biopsy	549	202	347	178	17	88.1 (82.84 to 92.24)	95.1 (92.27 to 97.12)
With biopsy	273	218	55	197	19	90.4 (85.65 to 93.94)	65.5 (51.42 to 77.76)

A patient with probability above the cut-off value (55%) specified by the EULAR/ACR classification criteria was defined as a positive case.
ACR, American College of Rheumatology; EULAR, European League Against Rheumatism; IIM, idiopathic inflammatory myopathies.

biopsy data was 0.97, 0.87 and 0.97, respectively (online supplementary figure 2). The specificity in the Japanese patients with muscle biopsy data being lower than that in the IMCCP cohort was addressed in the following separate analyses.

Sensitivity and specificity for subgroup classification of the Japanese cohort with the EULAR/ACR classification criteria and classification tree

The patients classified as having IIM according to the EULAR/ACR criteria were further classified into subgroups with the associated classification tree.[6 7] The percentages of the patients who were subclassified correctly with the classification tree into the 'true' classification made by their physicians were 70.9% in PM, 73.2% in DM, 82.9% in ADM, 100% in JDM and 71.4% in IBM (table 3). The subgroup classification performed well as in the IMCCP analysis.[6 7] Difficult cases included a 77-year-old patient, with serum CK activation but not with objective muscle weakness. The patient had erythematous papules on the extensor surfaces of extremities. The IMCCP definition of Gottron's papules is 'Erythematous to violaceous papules over the extensor surfaces of joints' and can be skin lesions of the other diseases including hand eczema and verruca. Thus, due to absence of any other cutaneous manifestations, the eruption of the patient was not judged as skin manifestation specific to IIM by a clinician. The patient was diagnosed as having PM, which was validated by experts, but fell into ADM according to the tree.

Comparison of sensitivity and specificity of the EULAR/ACR criteria with other classification criteria

Sensitivity and specificity of the new classification criteria with cut-off probability 55% in Japanese patients with PM/DM/JDM/ADM were 87.4% and 92.4%, respectively (table 4). They were then compared with previously published criteria (Bohan & Peter's and Tanimoto's criteria) using the same dataset (table 4). Since Bohan & Peter's and Tanimoto's criteria are for classification of PM/DM/JDM/other juvenile myositis, but not of ADM, the ADM cases were excluded in their specificity and sensitivity studies: actually only 26.3% (20 of 76) of patients with ADM were captured using Bohan and Peter's criteria. The specificity and specificity of the new criteria were comparable with those of Bohan & Peter's criteria (88.4% and 88.3%, respectively) and were higher than those of Tanimoto's criteria (82.2% and 87.8%, respectively).

In Japan, Tanimoto's criteria were partly revised in 2015 to include ADM by adding the following note to the criteria; an amyopathic patient whose histopathological findings of the eruptions are compatible with DM can be diagnosed as having ADM (online supplementary table 1). This change was made for expansion of medical expense subsidisation for patients with ADM. When patients with ADM were included for analyses of the revised Tanimoto's criteria, the sensitivity and specificity turned out to be 71.2% and 87.8%, respectively. Thus, the EULAR/ACR classification criteria have the highest sensitivity and/or specificity for the Japanese cohort, regardless of whether ADM is included or not.

Sensitivity and specificity of the EULAR/ACR classification criteria for IIM after overlap syndrome were included in IIM

The specificity for patients with the muscle biopsy data was low (65.5%) (table 2) in the Japanese cohort. This was because 19 out of 55 patients with non-IIM with muscle biopsy data were diagnosed as having IIM according to the new criteria. The diagnosis of the 19 cases included rimmed vacuolar distal myopathy (all three patients were misclassified), mitochondrial myopathy (1 of 6 patients), sarcoidosis (2 of 3 patients), systemic vasculitis (1 of 11 patients), other dystrophy (4 of 10 patients), systemic

Table 3 Comparison of physician-diagnosed IIM subgroups with IIM subgroups defined according to the classification tree among patients meeting the EULAR/ACR classification criteria for IIM

	True subgroup in IIM						
Classified subgroup*	PM	DM	ADM	JDM	IMNM	IBM	Total
Not IIM (<55% probability)	29	9	6	0	0	1	45
PM (IMNM)	95	14	0	0	2	1	112
DM	3	142	6	0	1	0	152
ADM	1	28	63	0	0	0	92
JDM	0	0	0	6	0	0	6
IBM	6	0	0	0	0	5	11
Unknown subgroup	0	1	1	0	0	0	2
Total	134	194	76	6	3	7	420
Correctly classified (%)	70.9	73.2	82.9	100.0	67.0	71.4	

*Classified subgroup by the EULAR/ACR classification criteria and classification tree, cut-off probability=55%.
ACR, American College of Rheumatology; ADM, amyopathic DM; DM, dermatomyositis; EULAR, European League Against Rheumatism; IBM, inclusion body myositis; IIM, idiopathic inflammatory myopathies; IMNM, immune-mediated necrotising myopathy; JDM, juvenile DM; PM, polymyositis.

Table 4 Sensitivity and specificity of EULAR/ACR classification criteria and the classification tree for PM/DM

	No. of subjects			No. of positive cases		Sensitivity	Specificity
	Total	PM/ DM	Non-PM/DM	PM/DM	Non-PM/DM	% (95% CI)	% (95% CI)
Classification criteria and tree							
Total	822	413	409	361	31	87.4 (83.82 to 90.45)	92.4 (89.41 to 94.79)
Without biopsy	549	202	347	177	17	87.6 (82.27 to 91.83)	95.1 (92.27 to 97.12)
With biopsy	273	211	62	184	14	87.2 (81.93 to 91.40)	77.4 (65.03 to 87.07)
	No. of subjects			No. of positive cases		Sensitivity	Specificity
	Total	PM/DM	Non-PM/DM	PM/DM	Non-PM/DM	% (95% CI)	% (95% CI)
Bohan & Petera*	746	337	409	298	48	88.4 (84.52 to 91.64)	88.3 (84.74 to 91.22)
Tanimotoa*	746	337	409	277	50	82.2 (77.69 to 86.13)	87.8 (84.20 to 90.79)
Revised Tanimotob†	822	413	409	294	50	71.2 (66.56 to 75.51)	87.8 (84.20 to 90.79)

A patient with probability above a specified cut-off value (55%) by the EULAR/ACR classification criteria and classified into PM/DM (including ADM, JDM and IMNM) by the classification tree was defined as a positive case.
A patient who fulfilled the Tanimoto's original or revised criteria or classified as having probable PM/DM in the Bohan & Peter's criteria was defined as a positive case.
*Analysis of PM/DM/JDM cases.
†Analysis of all PM/DM/JDM/ADM cases.
ACR, American College of Rheumatology; ADM, amyopathic DM; DM, dermatomyositis; EULAR, European League Against Rheumatism; IBM, inclusion body myositis; IIM, idiopathic inflammatory myopathies; IMNM, immune-mediated necrotising myopathy; JDM, juvenile DM; PM, polymyositis.

sclerosis (2 of 3 patients) and mixed connective tissue disease (MCTD, 6 of 8 patients). We suggest that Japanese physicians have differing definition of myositis from those from other countries when patients have other autoimmune diseases. When patients have myositis with systemic sclerosis or MCTD, they are diagnosed as having scleroderma myopathy or MCTD itself, but not IIM. In contrast, in the EULAR/ACR criteria, this is considered as overlap syndrome.

Suspecting that the low specificity may be due to the differing definition of myositis, we changed the definition. When patients with 'overlap syndrome' were defined as those with SLE, MCTD and systemic sclerosis, who had abnormalities in any of four items of muscle findings or elevation of serum skeletal muscle enzymes, 12 patients (1 SLE, 8 MCTD and 3 systemic sclerosis) with muscle biopsy data were diagnosed as having overlap syndrome. If these patients were transferred to IIM, the specificity in classification of the patients with biopsy increased up to 74.4% (table 5).

DISCUSSION

The present analysis, involving Japanese 420 IIM cases and 402 non-IIM comparator cases, is important as the first external validation of specificity of the new classification criteria. The sensitivity was also validated within a cohort of Asian ethnicity. There was some question as to whether the new criteria could be validated in Asian and African populations, since these groups were underrepresented in the IMCCP study.[6 7]

Another significance of this study is the comparison of the new criteria with the previous criteria. Bohan & Peter's criteria and Tanimoto's original or revised criteria have been commonly used in Japan. They were not designed for classification of IBM. Thus, we calculated the sensitivity and specificity of Tanimoto's criteria and Bohan & Peter's criteria for our Japanese patients and directly compared the usefulness of the new criteria and previous criteria for the classification of PM/DM/JDM/ADM. The specificity and/or sensitivity in the new criteria was higher than that of Bohan & Peter's criteria and Tanimoto's original or revised criteria, confirming that the new criteria should be applied in classification of Japanese patients with IIM instead of Tanimoto's revised criteria.

The low specificity of the new criteria for Japanese patients was due to myositis associated with other autoimmune diseases not being defined as PM in overlap syndromes. This is in sharp contrast with the classification in the EULAR/ACR areas, where these conditions are defined as overlap syndromes. Transfer of patients with those conditions from the comparator group to the IIM group actually made the specificity higher. To apply the new criteria in Japan, it appears that Japanese physicians should use the same definitions as those in the EULAR/ACR countries.

Other different points between the IMCCP and Japanese cohorts include the composition of diseases in non-IIM. The Japanese cohort had more rheumatic diseases (74% vs 36%), more dermatological diseases (15% vs 5%) and less neurological diseases (9% vs 58%) than the IMCCP cohort. The new criteria may be therefore better in ruling out neurological comparators than rheumatological and dermatological comparators when muscle biopsy was performed. Notably, the report of the EULAR/ACR criteria also described the sensitivity and specificity of Bohan & Peter's criteria (98% and 55%, respectively) and Tanimoto's original criteria (96% and 31%, respectively) for the

Table 5 Sensitivity and specificity of EULAR/ACR classification criteria for IIM after overlap syndrome was included in IIM

	No. of subjects			No. of positive cases		Sensitivity	Specificity
	Total	IIM	Non-IIM	IIM	Non-IIM	% (95% CI)	% (95% CI)
EULAR/ACR classification criteria for IIM							
Total	822	432	390	383	28	88.7 (85.28 to 91.49)	92.8 (89.79 to 95.18)
Without biopsy	549	202	347	178	17	88.1 (82.84 to 92.24)	95.1 (92.27 to 97.12)
With biopsy	273	230	43	205	11	89.1 (84.37 to 92.84)	74.4 (58.83 to 86.48)

A patient with probability above a specified cut-off value (55%) by the EULAR/ACR classification criteria was defined as a positive case.
ACR, American College of Rheumatology; EULAR, European League Against Rheumatism; IIM, idiopathic inflammatory myopathies.

classification of IIM in their cohort and concluded that the new criteria are superior in specificity. Low specificity of the previous criteria in the IMCCP analysis may be affected by the percentage of neurological comparators. Similarly, in the current project, very few children were enrolled in both IIM cases and comparators; thus, the results may mainly refer to the adult population.

Second, patients with IBM were underrepresented in the Japanese cohort (1.7% vs 18%). Fewer patients with IBM were enrolled in this study because we have less patients with IBM in Japan. The national registry of intractable diseases, which should cover all patients requiring subsidisation of medical expenses, had only 350 patients with IBM registered among the Japanese population of 127 million people in 2017. The new criteria might identify IBM more readily than other IIM.

Third, muscle biopsies were performed in IMCCP cohort mostly if skin rashes were absent. Thus, muscle biopsy data were available in 80% of cases and comparators.[6 7] In Japan, however, muscle biopsy is generally performed only when differential diagnosis is difficult. There may therefore be a greater number of difficult cases in the patient group with muscle biopsy data. The low number of individuals with available muscle biopsy data particularly in the comparator cases (only approximately 15%) may affect the statistics and the validity of the results. The present study is retrospective, so we could not control the number of patients with muscle biopsy data.

The present study has some limitations, as described in online supplementary text. Moreover, the new criteria did not include recently identified myositis-specific autoantibodies such as anti-MDA5, anti-Mi-2 and TIF1γ autoantibodies. In our analysis, several patients with DM could not be diagnosed correctly by the new criteria, but were positive for anti-ARS antibodies, which indicates that the new criteria could be improved by adding these specific antibodies. As was discussed in the IMCCP report together with the MRI and electromyogram findings,[6] contribution of these new antibodies to sensitivity and specificity should be validated in the future.

Author affiliations
[1]Department of Dermatology and Plastic Surgery, Faculty of Life Sciences, Kumamoto University, Kumamoto, Japan
[2]Division of Public Health, Department of Social Medicine, Faculty of Medicine, Saitama Medical University, Saitama, Japan
[3]Department of Neurology and Neurological Science, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan
[4]Department of Internal Medicine and Rheumatology, Juntendo University School of Medicine, Tokyo, Japan
[5]Department of Rheumatology, Endocrinology and Nephrology, Faculty of Medicine and Graduate School of Medicine, Hokkaido University, Hokkaido, Japan
[6]Department of Dermatology, University of Tsukuba, Ibaraki, Japan
[7]Department of Neurology and Clinical Neuroscience, Yamaguchi University Graduate School of Medicine, Yamaguchi, Japan
[8]Institute of Rheumatology, Tokyo Women's Medical University, Tokyo, Japan
[9]Department of Immunology and Rheumatology, Unit of Advanced Preventive Medical Sciences, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan
[10]Division of Rheumatic Diseases, National Center for Global Health and Medicine, Tokyo, Japan
[11]Department of Rheumatology and Clinical Immunology, Graduate School of Medicine, Kyoto University, Kyoto, Japan
[12]Department of Rheumatology and Applied Immunology, Faculty of Medicine, Saitama Medical University, Saitama, Japan
[13]Department of Dermatology, Nagoya University Graduate School of Medicine, Aichi, Japan
[14]Department of Rheumatology, Hyogo College of Medicine, Hyogo, Japan
[15]Department of Neurology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan
[16]Division of Rheumatology, Department of Internal Medicine, Keio University School of Medicine, Tokyo, Japan
[17]The First Department of Internal Medicine, School of Medicine, University of Occupational and Environmental Health, Fukuoka, Japan
[18]Department of Allergy and Rheumatology, The University of Tokyo Hospital, Tokyo, Japan
[19]Department of Internal Medicine, Faculty of Medicine, University of Tsukuba, Ibaraki, Japan
[20]Department of Rheumatology, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan

Acknowledgements This study was supported in part by project research on intractable diseases from the Japanese Ministry of Health, Labour and Welfare.

Contributors MJ, AO, SI and HK performed the research, analysed the data and wrote the paper. Other authors contributed to collect patients' data. MJ and AO equally contributed to this work.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement statement The clinical data of patients were retrospectively collected from the medical records using questionnaires.

Patient consent for publication Not required.

Ethics approval The project was approved by research ethics committees of individual institutes (No. 1580).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Author note Authors' affiliations listed were current at the time that the data collection was performed. The current address of Hajime Sano is: Department of Rheumatology, Kyoto Okamoto Memorial Hospital, Kyoto, Japan.

ORCID iDs
Masatoshi Jinnin http://orcid.org/0000-0001-9230-6207
Yoshinao Muro http://orcid.org/0000-0003-2329-8375
Yoshiya Tanaka http://orcid.org/0000-0002-0807-7139
Kazuhiko Yamamoto http://orcid.org/0000-0001-9037-3625

REFERENCES

1 Plotz PH, Rider LG, Targoff IN, *et al*. Nih conference. myositis: immunologic contributions to understanding cause, pathogenesis, and therapy. *Ann Intern Med* 1995;122:715–24.
2 Medsger TA, Dawson WN, Masi AT. The epidemiology of polymyositis. *Am J Med* 1970;48:715–23.
3 Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). *N Engl J Med* 1975;292:344–7.
4 Bohan A, Peter JB. Polymyositis and dermatomyositis (second of two parts). *N Engl J Med* 1975;292:403–7.
5 Tanimoto K, Nakano K, Kano S, *et al*. Classification criteria for polymyositis and dermatomyositis. *J Rheumatol* 1995;22:668–74.
6 Lundberg I, Tjärnlund A, Bottai M, *et al*. European League against Rheumatism/American College of rheumatology classification criteria for adult and juvenile idiopathic inflammatory myopathies and their major subgroups. *Ann Rheumatol Dis* 2017;2017:2271–82.
7 Bottai M, Tjärnlund A, Santoni G, *et al*. EULAR/ACR classification criteria for adult and juvenile idiopathic inflammatory myopathies and their major subgroups: a methodology report. *RMD Open* 2017;3:e000507.

CLINICAL SCIENCE

Novel ultrasonographic Halo Score for giant cell arteritis: assessment of diagnostic accuracy and association with ocular ischaemia

Kornelis S M van der Geest ,[1,2] Frances Borg,[2] Abdul Kayani,[2] Davy Paap,[1,3] Prisca Gondo,[2] Wolfgang Schmidt,[4] Raashid Ahmed Luqmani,[5] Bhaskar Dasgupta [2]

Handling editor Josef S Smolen

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-216343).

For numbered affiliations see end of article.

Correspondence to
Dr Bhaskar Dasgupta, Rheumatology, Southend University Hospital NHS Foundation Trust, Westcliff-on-Sea, UK;
bhaskar.dasgupta@southend.nhs.uk

Received 19 September 2019
Revised 18 November 2019
Accepted 16 December 2019
Published Online First 3 January 2020

Watch Video
ard.bmj.com

ABSTRACT

Objectives Ultrasound of temporal and axillary arteries may reveal vessel wall inflammation in patients with giant cell arteritis (GCA). We developed a ultrasound scoring system to quantify the extent of vascular inflammation and investigated its diagnostic accuracy and association with clinical factors in GCA.

Methods This is a prospective study including 89 patients suspected of having GCA, of whom 58 had a confirmed clinical diagnosis of GCA after 6 months follow-up. All patients underwent bilateral ultrasound examination of the three temporal artery (TA) segments and axillary arteries, prior to TA biopsy. The extent of vascular inflammation was quantified by (1) counting the number of TA segments and axillary arteries with a halo and (2) calculating a composite Halo Score that also incorporated the thickness of each halo.

Results Halo counts and Halo Scores showed moderate diagnostic accuracy for a clinical diagnosis of GCA. They correlated positively with systemic inflammation. When compared with the halo count, the Halo Score correlated better with C-reactive protein (CRP) levels and allowed to firmly establish the diagnosis of GCA in more patients. Higher halo counts and Halo Scores were associated with a higher risk of ocular ischaemia. They allowed to identify subgroups of patients with low risk (≤5%) and high risk of ocular ischaemia (>30%).

Conclusions Ultrasound halo scoring allows to quantify the extent of vascular inflammation in GCA. Extensive vascular inflammation on ultrasound may provide strong diagnostic confirmation and associates with ocular ischaemia in GCA.

Key messages

What is already known about this subject?

- Vascular ultrasound is the recommended first-line investigation in the diagnostic work-up of giant cell arteritis (GCA).

What does this study add?

- This prospective study shows that the extent of vascular inflammation on ultrasound, that is, a halo count ≥2 or Halo Score ≥3, identifies GCA patients at high risk for ocular ischaemia (>30%). Patients below these cut-off points have a low risk for ischaemic vision loss (≤5%). Furthermore, halo counts and Halo Scores associate with systemic inflammation.
- The presence of a high Halo Score ≥10 may help to firmly diagnose GCA in a substantial portion of patients, that is, with high specificity (>95%) and a high positive likelihood ratio (>5.0).

How might this impact on clinical practice or future developments?

- Quantifying the extent of vascular inflammation by ultrasound has diagnostic value and may help to discriminate between patients with a high or low risk for ocular ischaemia.
- The halo count and Halo Score await further validation and could potentially be interesting outcome parameters in translational and therapeutic studies.

INTRODUCTION

Giant cell arteritis (GCA) is an autoimmune disease characterised by inflammation of large-sized and medium-sized arteries. Ocular ischaemia is a feared complication of GCA.[1] Laboratory testing often reveals systemic inflammation, that is, high C-reactive protein (CRP) levels, anaemia and thrombocytosis.[2]

EULAR recommendations identify temporal and axillary artery ultrasound as the first-line investigation in patients suspected of having GCA.[3] A halo is the main ultrasound finding suggestive of GCA.[4 5] A halo is a homogeneous, hypoechoic wall thickening of the artery, reflecting inflammation-induced oedema of the arterial wall.[6] Ultrasound has a 77% sensitivity and 96% specificity for GCA.[7]

Little is known about the relationship between the extent of vascular inflammation on ultrasound and disease severity in GCA. Aschwanden *et al* evaluated 11 vascular regions for the presence of a halo and showed that involvement of large systemic arteries is associated with more weight loss.[8] Schmidt *et al* linked axillary artery involvement to a low risk of eye complications.[9] The risk of eye complications was not related to the number of temporal artery (TA) segments with a halo in the latter study. In neither of these studies was halo thickness incorporated into the analysis of disease extent.

In the current study, we evaluated whether the extent of vascular inflammation on ultrasound could be linked to disease severity in GCA. To enumerate disease extent, we first calculated the number of TA segments and axillary arteries with

Check for updates

To cite: van der Geest KSM, Borg F, Kayani A, *et al*. *Ann Rheum Dis* 2020;**79**:393–399.

a halo sign. Furthermore, we developed a novel Halo Score that encompassed both the number of halos and the maximum halo thickness in each vascular region. We investigated the diagnostic accuracy of halo counts and Halo Scores, and their relationship with disease severity as indicated by ocular ischaemia and systemic inflammation.

METHODS

Patients

Consecutive patients suspected of having GCA (n=104) were prospectively recruited at the Rheumatology Department of Southend University Hospital between June 2010 and December 2013 as part of the TABUL study.[10] Patients underwent arterial ultrasound followed by a TA biopsy (TAB). Ultrasound and TAB were performed within 7 days after initiation of high-dose glucocorticoids. Patients were re-assessed after 6 months. The reference standard for GCA was the final clinical diagnosis after 6 months (online supplementary methods).

Ultrasound

Ultrasound scans were performed by a single, experienced ultrasonographer (BD) with an Esaote MyLab70 or MyLabTwice. A linear probe (LA435) with a grey-scale frequency of 18 MHz and colour Doppler frequency of 9 MHz was used. The focus was positioned at 5 mm below the skin for the TA. The pulse repetition frequency was 2–3 kHz. The colour box was set at an angle of at least 60°. The gain setting was adjusted to just fill the lumen. Patients were lying in a (semi-)recumbent position during the examination. The common superficial TA, its parietal and frontal branches, as well as the axillary arteries were fully and bilaterally examined in the long and short planes. In each vascular territory, the thickness of the largest halo was measured with one decimal place at the point of maximum thickness in the longitudinal plane. The ultrasonographer was not blinded to the clinical data of the patient. An ultrasound expert panel evaluated all scans and reports to monitor the scan quality and the adequacy of the reported findings. A halo sign was morphologically defined as an ultrasound finding of a dark hypoechoic area around the vessel lumen. A composite Halo Score was developed based on percentiles of halo thickness in patients with GCA.

Statistics

Information regarding statistics is provided in the figure legends and online supplementary methods. P values <0.05 were considered statistically significant. Data were analysed with IBM SPSS Statistics V.25, StatsDirect V.3.1.22 and Graphpad Prism V.5.

RESULTS

Patient characteristics

Out of 104 patients with suspected GCA, 92 patients underwent both ultrasound and TAB at baseline, and 89 patients completed 6 months follow-up (table 1). A clinical diagnosis of GCA was established in 58 out of 89 patients. Diagnoses in non-GCA patients are shown in online supplementary table S1.

Halo thickness and construction of the Halo Score

At baseline, the three TA segments and the axillary artery were examined by ultrasound on each side. In GCA patients, halos were reported in 41 common TA segments, 29 parietal TA segments, 32 frontal TA segments and 14 axillary arteries (figure 1A). If present, the maximum thickness of the halo was measured (figure 1B).

In order to develop a composite Halo Score, we identified the 20%, 40%, 60% and 80% percentiles of the maximum halo thickness in the TA segments and axillary arteries of patients with GCA (online supplementary table S2). Based on these arbitrary percentiles, we assigned halo grade scores to each TA segment and axillary artery (figure 1C). The distribution of halo grades among GCA patients with a halo is shown in online supplementary tables S3 and S4.

The sum of all halo grade scores was used to construct the Halo Score for each patient (figure 1D). To give equal weight to temporal and axillary arteries, the halo grade scores of the axillary arteries were multiplied by a factor of 3. Therefore, the Halo Score values could range from 0 to 48. For halo counts in TA segments and axillary arteries, no correction factor was used for the axillary artery. Thus, halo counts could vary from 0 to 8.

Diagnostic accuracy for GCA

Baseline halo counts and Halo Scores were higher in patients with a subsequently confirmed diagnosis of GCA than patients without GCA (figure 2A,B). Two non-GCA patients showed a high halo count. These halos were small in one male patient and could be attributed to atherosclerosis in one female patient (online supplementary table S5). Halo counts and Halo Scores showed similar diagnostic accuracy for a clinical diagnosis of GCA, as indicated by an area under the curve (AUC) of >0.70 in the receiver operating characteristic (ROC) curve analysis (figure 2C,D). At the optimal cut-off point, the sensitivity/specificity and likelihood ratios were comparable for both ultrasound parameters. In a subanalysis restricted to halo counts/Halo Scores in the TA only, similar diagnostic accuracy was obtained (online supplementary table S6).

Alternative cut-off points providing a specificity of 95% for a clinical diagnosis of GCA could be obtained: a halo count of ≥6, or Halo Score of ≥10. Although a Halo Score of ≥10 was present in 12 patients (21% of all patients with GCA), only two patients (3% of all patients with GCA) showed a halo count of ≥6 (online supplementary table S7). The positive likelihood ratio of a Halo Score ≥10 was high (LR +6.41), but poor for the halo count at this cut-off point (LR +1.07). Thus, the Halo Score could be used more effectively than the halo count to establish a diagnosis of GCA in more patients.

Diagnostic accuracy for positive TAB

The frequency and thickness of halos was higher in GCA patients with a positive TAB than patients with a negative TAB (online supplementary table S8). Consequently, halo counts and Halo Scores were higher in patients with a positive TAB than those with a negative biopsy (figure 2E,F). Both ultrasound parameters showed a good ability to predict a positive TABwith an AUC >0.80 in the ROC analysis (figure 2G,H). The sensitivity and specificity, positive likelihood ratios>2 and negative likelihood ratios<0.5 indicated that halo counts and Halo Score may help to predict the TAB result. Comparable diagnostic accuracy was obtained, if only TAs halo counts were taken into account (online supplementary table S6).

Effect of glucocorticoid treatment

Halo signs may disappear within days to weeks following initiation of glucocorticoid treatment.[5 11] When comparing patients with GCA receiving glucocorticoids for 0–1 days, 2–3 days and 4–7 days prior to ultrasound, we did not observe any differences in halo counts or Halo Scores (online supplementary figure S1). Patients using glucocorticoids for 4–7 days prior to ultrasound

Table 1 Patients' characteristics

Patients' characteristics	All patients (n=89)	Patients with GCA (n=58)	Patients without GCA (n=31)
Sex, no. of males	26 (29%)	15 (26%)	11 (36%)
Age, median (range) years	73 (44–96)	74 (50–96)	67 (44–90)
High-dose steroids started ≤7 days before baseline, no. of patients	75 (84%)	49 (85%)	26 (84%)
TAB positive according to pathologist, no. of patients	26 (29%)	26 (45%)	0 (0%)
TAB length, median (range) mm	7 (2–20)	7 (2–20)	8 (2–13)
Fulfilling 1990 ACR criteria for GCA, no of patients	72 (81%)	50 (86%)	22 (71%)
Any head pain present, no of patients	85 (96%)	55 (95%)	30 (97%)
New localised head pain, no of patients	77 (87%)	48 (83%)	29 (94%)
New generalised scalp tenderness, no of patients	52 (58%)	35 (60%)	17 (55%)
Swelling over temporal artery, no of patients	22 (25%)	14 (24%)	8 (26%)
Pain over temporal artery, no of patients	49 (55%)	29 (50%)	20 (65%)
Jaw claudication, no of patients	42 (47%)	32 (55%)	10 (32%)
Tongue claudication, no of patients	3 (3%)	2 (3%)	1 (3%)
Any visual symptoms, no of patients	47 (53%)	30 (52%)	17 (55%)
Reduced or lost vision, no of patients	38 (43%)	26 (45%)	12 (39%)
Double vision, no of patients	9 (10%)	4 (7%)	5 (16%)
Amaurosis fugax, no of patients	2 (2%)	2 (3%)	0 (0%)
Anorexia, no of patients	31 (35%)	22 (38%)	9 (29%)
Fatigue, no of patients	65 (73%)	42 (72%)	23 (74%)
Fever or night sweats, no of patients	38 (43%)	25 (43%)	13 (42%)
Polymyalgia, no of patients	16 (18%)	14 (24%)	2 (7%)
Temporal artery thickening, no of patients	28 (32%)	21 (36%)	7 (23%)
Temporal artery tenderness, no of patients	50 (56%)	29 (50%)	21 (68%)
Temporal artery abnormal pulse, no of patients	18 (20%)	16 (28%)	2 (7%)
Axillary artery tenderness, no of patients	8 (9%)	5 (9%)	3 (10%)
AION*, no of patients	15 (17%)	10 (17%)	5 (16%)
PION*, no of patients	5 (6%)	2 (3%)	3 (10%)
RAPD*, no of patients	7 (8%)	5 (9%)	2 (7%)
Ocular ischaemia (AION/PION/RAPD), no of patients	19 (21%)	12 (21%)	7 (23%)
Ocular palsy*†, no of patients	0 (0%)	0 (0%)	0 (0%)
Bruits*, no of patients	0 (0%)	0 (0%)	0 (0%)
Stroke*	2 (2%)	0 (0%)	2 (7%)
ESR, mm/hour,† median (range)	34 (3–90)	44 (3–90)	9 (3–77)
CRP, mg/L,† median (range)	46 (3–329)	54 (3–329)	13 (3–205)
Haemoglobin (g/dL), median (range)	12.8 (8.9–16.0)	12.0 (8.9–15.5)	13.5 (10.1–16.0)
Platelets, 10^9/L, median (range)	343 (126–661)	363 (167–661)	317 (126–522)

Details of the 89 patients recruited in the TABUL study at Southend University Hospital, who underwent ultrasound, temporal artery biopsy and 6 months follow-up.
ESR was determined in n=57 patients and CRP in n=54 subjects. ESR and CRP were measured before initiation of high-dose glucocorticoid treatment. Haemoglobin levels and platelet counts were determined prior to high-dose glucocorticoid treatment or within 7 days after initiation of this treatment.
*Considered negative if not reported.
†ESR and CRP were not performed in every subject.
AION, anterior ischaemic optic neuropathy;CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; GCA, giant cell arteritis; PION, posterior ischaemic optic neuropathy; RAPD, relative afferent pupillary defect; TAB, temporal artery biopsy.

tended to have a higher prevalence of ocular ischaemia and polymyalgic symptoms when compared with other patients, although not statistically significant (online supplementary table S9).

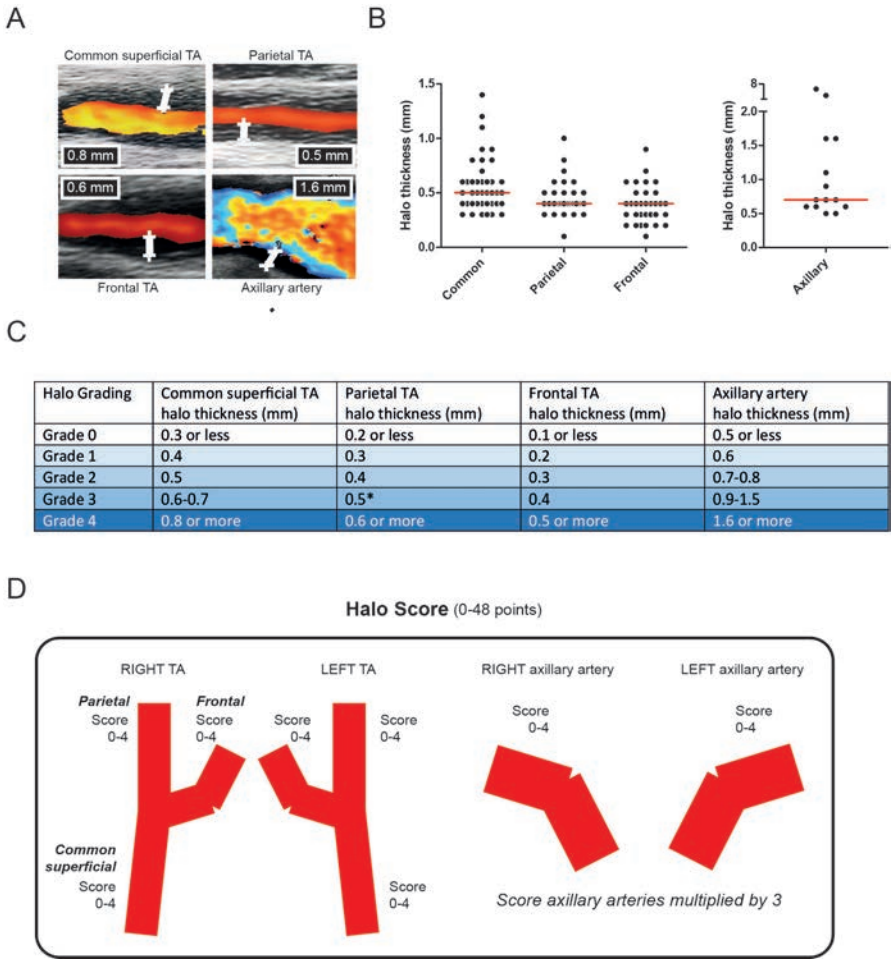
Vascular and systemic inflammation

We questioned if ultrasound findings could be linked to systemic inflammation in patients with GCA. Halo counts showed no correlation with haemoglobin levels but correlated positively with CRP levels and platelet counts (figure 3A). The Halo Score correlated even better with CRP levels, showed a positive correlation with platelets counts, and correlated negatively with haemoglobin levels (figure 3B). In a subanalysis of halo counts/Halo Scores restricted to the TA only, these correlations became less clear (online supplementary table S10). The presence of axillary artery involvement tended to be associated with more systemic inflammation (online supplementary table S11). Taken together, Halo Scores associated stronger with systemic inflammation than halo counts.

The erythrocyte sedimentation rate (ESR), as measured by a capillary photometric-kinetic technique (Alifax),[12] was remarkably low in patients with GCA (online supplementary table S12). Pretreatment ESR was <30 mm/hour in 31% of GCA patients, while CRP was <10 mg/L in 2% of patients. Only 46% of GCA patients showed an ESR >50 mm/hour. Thus, the ESR showed no correlation with CRP levels, halo counts or Halo Scores (online supplementary figure S2).

Extent of vascular inflammation and patients' characteristics

We performed multiple linear regression analysis to investigate if ocular ischaemia, or perhaps other clinical characteristics, were associated with higher halo counts or Halo Scores. Ocular ischaemia was defined by the presence of anterior ischaemic optic neuropathy (AION), posterior ischaemic optic neuropathy (PION) and/or a relative afferent pupillary defect (RAPD). Both ocular ischaemia and male sex were independently associated with higher halo counts and Halo Scores in patients with GCA



Halo Grading	Common superficial TA halo thickness (mm)	Parietal TA halo thickness (mm)	Frontal TA halo thickness (mm)	Axillary artery halo thickness (mm)
Grade 0	0.3 or less	0.2 or less	0.1 or less	0.5 or less
Grade 1	0.4	0.3	0.2	0.6
Grade 2	0.5	0.4	0.3	0.7-0.8
Grade 3	0.6-0.7	0.5*	0.4	0.9-1.5
Grade 4	0.8 or more	0.6 or more	0.5 or more	1.6 or more

Figure 1 Ultrasound halo scoring. (A) Representative ultrasound images of halo signs, and measurements of halo thickness, in the common superficial TA, parietal TA, frontal TA and axillary artery. (B) Thickness of halo signs that were reported in 41 common TA segments, 29 parietal TA segments, 32 frontal TA segments and 14 axillary arteries of patients with GCA. (C) Halo grade scoring system and cut-off values. Due to similar 40% and 60% percentile boundaries in the parietal TA, a cut-off value of 0.5 mm was used for a grade 3 halo in this TA segment. (D) Construction of the Halo Score. Axillary artery scores were multiplied by 3 to give equal weight to the TA and axillary artery for the Halo Score. TA, temporal artery.

(table 2). No further relationships were observed between clinical features and ultrasound parameters.

Diagnostic accuracy for ocular ischaemia

Halo counts and Halo Scores showed fair ability to discriminate between GCA patients with and without ocular ischaemia, as indicated by an AUC >0.70 in the ROC analysis (figure 4A). At the optimal cut-off point for ocular ischaemia, that is, halo count ≥2 or Halo Score ≥3, an excellent sensitivity and poor specificity were obtained. At the optimal cut-off points, positive likelihood ratios <2 indicated that halo counts and Halo Scores were not helpful in predicting the presence of ocular ischaemia. However, negative likelihood ratios were <0.2. Thus, low halo counts and Halo Scores helped to identify a substantial group of patients with a low risk of ocular ischaemia (figure 4B). In a subanalysis of halo counts and Halo Scores restricted to the TA only, comparable diagnostic accuracy for ocular ischaemia was obtained (online supplementary figure S3). The presence of axillary involvement per se showed no effect on the risk of ocular ischaemia (online supplementary table S13).

Next, we evaluated if halo counts and Halo Scores were independent predictors for ocular ischaemia in a logistic regression analysis (online supplementary table S14). A halo count ≥2 provided an OR for ocular ischaemia of 12.000 (95% CI=1.430 to 100.705; p=0.022), whereas a Halo Score ≥3 showed an OR of 9.880 (95% CI=1.137 to 85.887; p=0.038). Other clinical characteristics were not predictive of ocular ischaemia. Thus, halo counts and Halo Scores were independent predictors of ocular ischaemia.

DISCUSSION

We show that the extent of vascular inflammation on ultrasound, as quantified by the halo count and novel Halo Score, can be linked to ocular ischaemia and systemic inflammation in GCA. The Halo Score allowed to firmly establish a diagnosis of GCA in more patients than the halo count.

The extent of inflammation was measured in the three TA segments and axillary arteries. Subclavian and facial arteries were not evaluated. However, axillary artery involvement identifies the vast majority of patients with inflammation of large systemic arteries,[13] whereas TA involvement identifies nearly all patients with cranial artery involvement.[14] EULAR recommendations recognise temporal and axillary artery ultrasound as the first-line investigation in GCA.[3] Examination of temporal and axillary arteries might therefore provide a reasonable estimation of disease extent in GCA.

Extensive vascular inflammation identified GCA patients at high (>30%) risk of ocular ischaemia. However, half of patients showed low halo counts/Halo Scores and a ≤5% risk

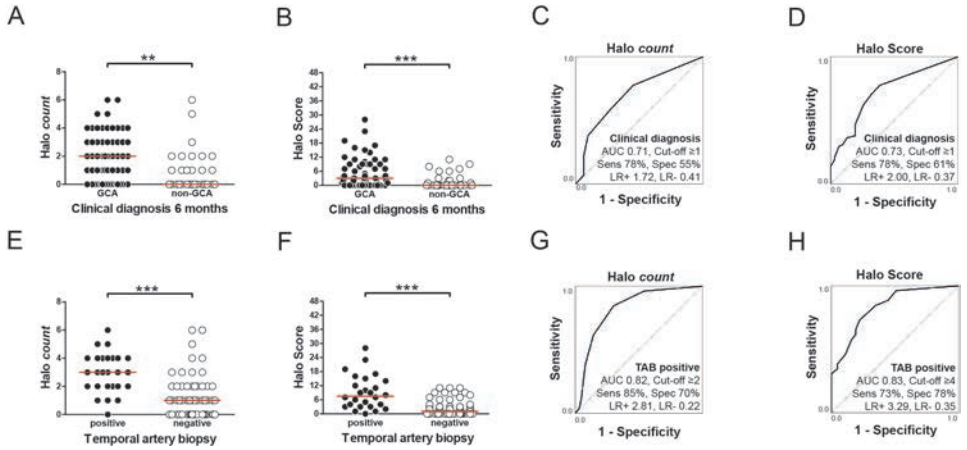


Figure 2 Diagnostic accuracy of halo count and Halo Score for GCA. (A) Baseline halo count in common superficial TAs, parietal TAs, frontal TAs and axillary arteries and (B) Halo Scores in patients with an eventually confirmed clinical diagnosis of GCA (n=58) versus non-GCA patients (n=31). (C) ROC curve showing the diagnostic accuracy of baseline halo counts and (D) Halo Scores for an eventual clinical diagnosis of GCA after 6 months. The optimal cut-off points were determined by Youden index. (E) Baseline halo counts and (F) Halo Scores in patients with a positive TAB (n=26) versus patients with a negative TAB (n=63). Overall, 25 TABs showed transmural inflammation and/or giant cells. One TAB considered positive for GCA showed an adventitial infiltrate, elastic lamina disruption and intimal hyperplasia without transmural inflammation/giant cells. (G) ROC curve showing the diagnostic accuracy of halo counts and (H) Halo Scores for a positive TAB. The optimal cut-off point was determined by Youden index. AUC, area under the curve; LR+, positive likelihood ratio; LR–, negative likelihood ratio; ROC, receiver operating characteristic; Sens, sensitivity; Spec, specificity; TA, temporal artery; TAB, temporal artery biopsy. Statistical significance at (A, B, E, F) was tested by Mann-Whitney U test: **p<0.01, ***p<0.001.

of GCA-related vision loss. As visual symptoms described by patients are not always related to GCA, we strictly defined ocular ischaemia by the presence of AION, PION and/or RAPD. Patients with suggestive eye symptoms were referred to the ophthalmologist in this single hospital study. Nevertheless, it might be a bias that not every patient underwent ophthalmological examination. Previously, no relationship was noted between the number of TA segments with a halo and ocular complications.[9] However, the definition of ocular complications in the latter study was broader than in the current study. Wall thickening of arteries supplying the retina is thought to cause ocular ischaemia in GCA.[1 15] Our findings indicate that wall thickening in the latter arteries likely parallels that in other arteries in GCA.

The extent of vascular inflammation correlated well with systemic inflammation in patients with GCA. Halo counts correlated positively with CRP levels and platelets counts. Halo Scores correlated even better with CRP levels than halo counts and also showed an inverse correlation with haemoglobin levels. No association was found with the ESR, which was measured by a capillary photometric–kinetic technique. This method provides an indirect estimation of the ESR[12] and might be less accurate than the traditional Westergren in the context of rheumatic inflammatory diseases.[16 17] Overall, our findings suggest a link between arterial and systemic inflammation in GCA.

Halo counts and Halo Scores showed comparable diagnostic accuracy for a clinical diagnosis of GCA. At the optimal cut-off

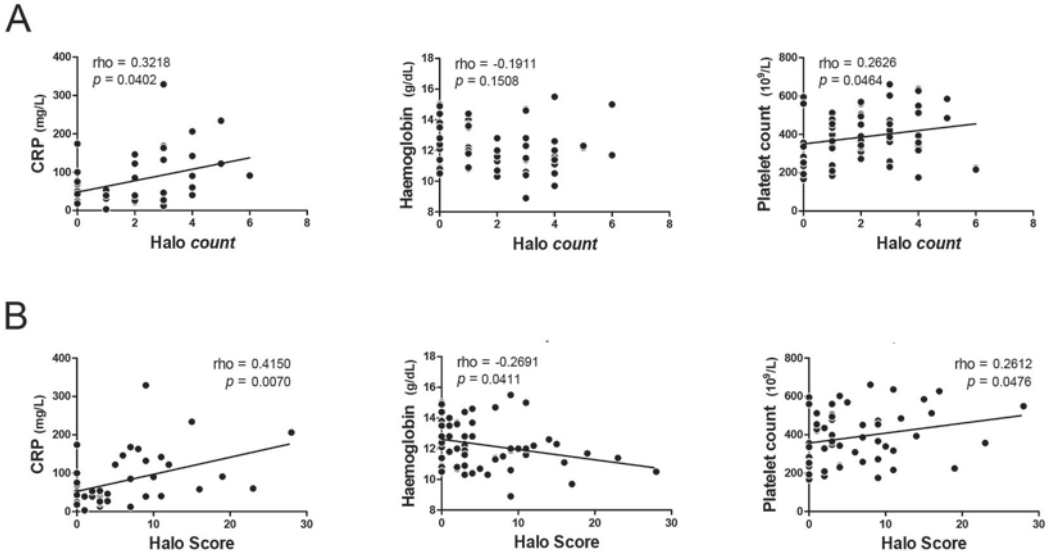


Figure 3 Relationship of halo count and Halo Score with systemic inflammation. (A) Correlation of halo counts and (B) Halo Scores with CRP, haemoglobin and platelets in patients with a clinical diagnosis of GCA. CRP levels were determined prior to initiation of treatment in 41 GCA patients. Haemoglobin levels and platelet counts were measured prior to treatment or within 7 days after initiation of high-dose glucocorticoids in 58 GCA patients. Correlations were determined by Spearman's rank correlation coefficient.

Table 2 Variables predicting the extent of vascular inflammation on ultrasound

Dependent variable	Predicting variable	Final model of multiple linear regression B (95% CI)	P value
Halo count	Age	–	
	Sex	1.109 (0.172 to 2.047)*	0.021
	Ocular ischaemia	1.103 (0.089 to 2.116)*	0.034
	Polymyalgia	–	
	Two or more systemic symptoms	–	
	Temporal artery palpable changes	–	
Halo Score	Age	–	
	Sex	2.902 (0.100 to 6.984)†	0.041
	Ocular ischaemia	3.488 (0.305 to 8.143)†	0.028
	Polymyalgia	2.813 (–0.053 to 7.080)†	0.056
	Two or more systemic symptoms	–	
	Temporal artery palpable changes	–	

Data are shown for baseline halo count and Halo Scores in patients with GCA (n=58). Multiple linear regression analysis was performed with backward exclusion of predicting variables. Since the Halo Score was not normally distributed, the Halo Score was transformed by square root. The probability of F for removal was 0.10. Results of the final model are shown. Age in years. Sex: 0=female, 1=male. Ocular ischaemia (ie, anterior ischaemic optic neuropathy, posterior ischaemic optic neuropathy and/or relative afferent pupillary defect), polymyalgia, two or more systemic symptoms (ie, anorexia, fever/night sweats, fatigue), temporal artery palpable changes (ie, thickening and/or loss of pulse): 0=absent, 1=present. (–) Variable removed due to backward exclusion.
*R^2=0.157, F(2,55) = 5.138, p=0.009.
†R^2=0.207, F(3,54) = 4.688, p=0.006.
GCA, giant cell arteritis.

points, both ultrasound parameters provided fair sensitivity (78%) but moderate specificity (55%–61%) for a diagnosis of GCA. Even better diagnostic accuracy was obtained for a positive TAB, which supports the idea that ultrasound might replace a TAB under certain conditions.[3] Alternative cut-off points providing 95% specificity for a clinical diagnosis of GCA could also be obtained. Few patients with GCA fulfilled this cut-off point for the halo count. In contrast,>20% of patients with GCA showed Halo Scores above the 95% specificity cut-off point, that is, a score ≥10. At this cut-off point, a high positive likelihood ratio could be obtained for the Halo Score, which allowed us to make a confident diagnosis of GCA in a substantial portion of patients.

Male sex was associated with higher halo counts and Halo Scores in patients with GCA. Recently, male sex predicted the presence of a halo sign on ultrasound in patients with GCA.[18] It might be possible that GCA is associated with more arterial thickening in men than women. However, it is also conceivable that the arterial calibre and arterial wall thickness are in general higher in men than women.[5] It would be interesting to collect sex-specific, normative data on arterial wall thickness.

A halo was morphologically defined as a dark hypoechoic area around the vessel lumen. As the halo compression sign was reported at the end of our study,[19] this sign was not tested. Recently, diagnostic cut-off values have been described for the intima–media thickness in TAs, as measured by a 22 MHz transducer.[20] Although still considered state of the art, our 18 MHz transducer frequently does not allow us to visualise the intima–media complex of the TAs. Halo thickness and intima–media thickness are therefore not fully interchangeable parameters. Halo counts and thickness might have been higher if measured with higher frequency transducers. In accordance with Monti *et al*,[21] we observed relatively similar halo thickness among the three TA segments. Halo counts in the latter study were comparable with those in the current study.

The same morphological halo definition was applied to the axillary arteries. The dark hypoechoic halo pattern differs from the normal intima–media complex, which can be readily identified as a double line in the axillary artery.[6] Part of halos reported in the axillary artery were smaller than a recently proposed diagnostic cut-off value, that is, 1.0 mm.[20] However, two provisional reports have suggested that axillary arteries may be inflamed despite a halo thickness <1.0 mm on ultrasound.[22 23] Since halos <1.0 mm might still relate to disease activity, these halos were included in the halo count/Halo Score. Overall, our main study findings were not compromised by axillary artery involvement, as indicated by our subanalyses restricted to the TA only.

We observed no clear association between short-term glucocorticoid treatment and extent of vascular inflammation on ultrasound. Serial ultrasound examinations before and after initiation of glucocorticoid treatment have shown that it takes weeks to months before the majority of TA halos disappear, while only few axillary artery halos disappear.[5 11 24] In our study, lack of treatment effect may have two explanations. First, treatment duration might have been too short. Second, patients taking glucocorticoids for 4–7 days showed a slightly higher prevalence of symptoms associated with higher halo counts/Halo Score than patients with shorter treatment duration.

A strong point of our study is its prospective design with patients undergoing ultrasound and TAB according to a fixed protocol.

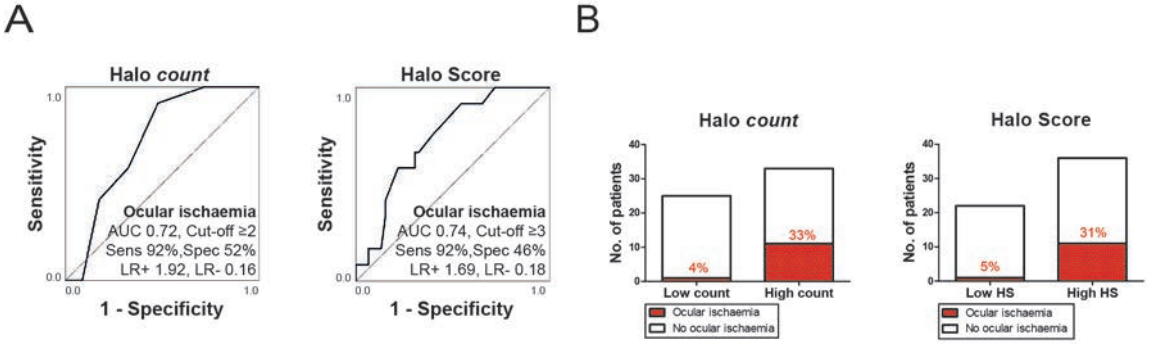


Figure 4 Halo count and Halo Score associated with ocular ischaemia. (A) Receiver operating characteristic curve showing diagnostic accuracy of baseline halo count (left panel) and Halo Score (right panel) for concomitant presence of ocular ischaemic symptoms. Ocular ischaemia was defined as the presence of anterior ischaemic optic neuropathy, posterior ischaemic optic neuropathy and/or a relative afferent pupillary defect. The optimal cut-off point was determined by Youden index. (B) Presence of ocular ischaemia (percentages are shown) among patients with low versus high halo count (left panel), or low versus high Halo Score (right panel) as determined by the optimal cut-off points mentioned at (A). AUC, area under the curve; HS, Halo Score; LR+, positive likelihood ratio; LR–, negative likelihood ratio; Sens, sensitivity; Spec, specificity.

The clinical diagnosis was rigorously established after 6 months follow-up. Ultrasound was performed by an experienced ultrasonographer. Our study has also potential limitations. There was a bias towards cranial GCA. The ultrasonographer was not blinded to the clinical data. However, a symptom likely to bias the ultrasonographer, that is, an abnormal TA on palpation, showed no effect on halo counts or Halo Scores. The intra-rater and inter-rater reliabilities were not tested, and should be evaluated in future studies. Nevertheless, the quality of the ultrasound scans and reports was monitored by an expert panel. Our findings were derived from a post-hoc analysis of a diagnostic trial and obtained from a single centre.[10] The Halo Score should be further validated by currently ongoing, prospective, multicentre studies (ClinicalTrials.gov NCT03765788; and NIHR Portfolio study #264294), prior to application in the clinic.

In conclusion, the extent of arterial inflammation in GCA can be quantified by ultrasound halo scoring. A high volume of vascular inflammation on ultrasound might strongly support the diagnosis of GCA and identifies patients at risk for ocular ischaemia. The clinical application of halo counts and Halo Scores warrants further validation in other studies.

Author affiliations
[1]Rheumatology and Clinical Immunology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
[2]Rheumatology, Southend University Hospital NHS Foundation Trust, Westcliff-on-Sea, UK
[3]Rehabilitation Medicine, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
[4]Medical Centre for Rheumatology Berlin-Buch, Immanuel-Krankenhaus GmbH, Berlin, Germany
[5]Rheumatology, University of Oxford Nuffield Department of Orthopaedics Rheumatology and Musculoskeletal Sciences, Oxford, UK

Twitter Bhaskar Dasgupta @profbdasgupta

Acknowledgements We thank all the patients and staff involved in the TABUL study.

Contributors KSMvdG, WS, RAL and BD contributed to conception or design of the work. FB, AK, PG, RAL and BD contributed to acquisition of data. KSMvdG, DP, WS, RAL and BD contributed to analysis or interpretation of data. All authors were involved in drafting of the work or revising it critically for important intellectual content. All authors provided final approval of the version published. All authors agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding The study was funded by the Health Technology Assessment Programme of the National Institute for Health Research (NIHR, #HTA-08/64/01).

Competing interests KSMvdG reports grants from the Mandema Stipend and the Dutch Society for Rheumatology, and personal fees from Roche, outside the submitted work. WS reports grants and personal fees from GSK, Novartis, Roche and Sanofi, and personal fees from Chugai, outside the submitted work. RAL reports grants from Arthritis UK, the Medical Research Council, the University of California San Francisco/Oxford Invention Fund, the Canadian Institutes of Health Research and The Vasculitis Foundation, grants and personal fees from GSK, and personal fees from Medpace, MedImmune and Roche, outside the submitted work. BD reports grants and personal fees from Roche, personal fees from GSK, BMS, Sanofi and AbbVie, outside the submitted work.

Patient and public involvement statement Advice on the study design was obtained from patients through PMRGCAuk. Patient representatives on the Trial Steering Committee and the Data Monitoring Committee provided important advice.

Patient consent for publication Not required.

Ethics approval The study was performed in accordance with the declaration of Helsinki. All patients provided written informed consent. The study was approved by the Berkshire Research Ethics Committee (REC#09/H0505/132).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iDs
Kornelis S M van der Geest http://orcid.org/0000-0003-2798-6765
Bhaskar Dasgupta http://orcid.org/0000-0002-5523-6534

REFERENCES

1 Patil P, Karia N, Jain S, *et al*. Giant cell arteritis: a review. *Eye Brain* 2013;5:23–33.
2 van der Geest KSM, Sandovici M, van Sleen Y, *et al*. Review: what is the current evidence for disease subsets in giant cell arteritis? *Arthritis Rheumatol* 2018;70:1366–76.
3 Dejaco C, Ramiro S, Duftner C, *et al*. EULAR recommendations for the use of imaging in large vessel vasculitis in clinical practice. *Ann Rheum Dis* 2018;77:636–43.
4 Schmidt W, Kraft H, Völker L, *et al*. Colour Doppler sonography to diagnose temporal arteritis. *The Lancet* 1995;345:866.
5 Schmidt WA, Kraft HE, Vorpahl K, *et al*. Color duplex ultrasonography in the diagnosis of temporal arteritis. *N Engl J Med* 1997;337:1336–42.
6 Chrysidis S, Duftner C, Dejaco C, *et al*. Definitions and reliability assessment of elementary ultrasound lesions in giant cell arteritis: a study from the OMERACT large vessel vasculitis ultrasound Working group. *RMD Open* 2018;4:e000598.
7 Duftner C, Dejaco C, Sepriano A, *et al*. Imaging in diagnosis, outcome prediction and monitoring of large vessel vasculitis: a systematic literature review and meta-analysis Informing the EULAR recommendations. *RMD Open* 2018;4:e000612.
8 Aschwanden M, Kesten F, Stern M, *et al*. Vascular involvement in patients with giant cell arteritis determined by duplex sonography of 2x11 arterial regions. *Ann Rheum Dis* 2010;69:1356–9.
9 Schmidt WA, Krause A, Schicke B, *et al*. Do temporal artery duplex ultrasound findings correlate with ophthalmic complications in giant cell arteritis? *Rheumatology* 2009;48:383–5.
10 Luqmani R, Lee E, Singh S, *et al*. The role of ultrasound compared to biopsy of temporal arteries in the diagnosis and treatment of giant cell arteritis (TABUL): a diagnostic accuracy and cost-effectiveness study. *Health Technol Assess* 2016;20:1–238.
11 De Miguel E, Roxo A, Castillo C, *et al*. The utility and sensitivity of colour Doppler ultrasound in monitoring changes in giant cell arteritis. *Clin Exp Rheumatol* 2012;30:S34–8.
12 Kratz A, Plebani M, Peng M, *et al*. ICSH recommendations for modified and alternate methods measuring the erythrocyte sedimentation rate. *Int J Lab Hematol* 2017;39:448–57.
13 Schmidt WA, Seifert A, Gromnica-Ihle E, *et al*. Ultrasound of proximal upper extremity arteries to increase the diagnostic yield in large-vessel giant cell arteritis. *Rheumatology* 2008;47:96–101.
14 Ješe R, Rotar Žiga, Tomšič M, *et al*. The role of colour Doppler ultrasonography of facial and occipital arteries in patients with giant cell arteritis: a prospective study. *Eur J Radiol* 2017;95:9–12.
15 Soriano A, Muratore F, Pipitone N, *et al*. Visual loss and other cranial ischaemic complications in giant cell arteritis. *Nat Rev Rheumatol* 2017;13:476–84.
16 van der Maas A, van den Ende CHM, van Eerd J, *et al*. The use of different methods for rapid determination of the ESR induces DAS28 misclassification in clinical practice. *Clin Exp Rheumatol* 2010;28:477–82.
17 Bernelot Moens HJ. [Temporal arteritis: do not rely on the erythrocyte sedimentation rate]. *Ned Tijdschr Geneeskd* 2015;159:A9196.
18 Conway R, O'Neill L, McCarthy GM, *et al*. Performance characteristics and predictors of temporal artery ultrasound for the diagnosis of giant cell arteritis in routine clinical practice in a prospective cohort. *Clin Exp Rheumatol* 2019;37 Suppl 117:72–8.
19 Aschwanden M, Daikeler T, Kesten F, *et al*. Temporal artery compression sign—a novel ultrasound finding for the diagnosis of giant cell arteritis. *Ultraschall Med* 2013;34:47–50.
20 Schäfer VS, Juche A, Ramiro S, *et al*. Ultrasound cut-off values for intima-media thickness of temporal, facial and axillary arteries in giant cell arteritis. *Rheumatology* 2017;56:1632.
21 Monti S, Floris A, Ponte CB, *et al*. The proposed role of ultrasound in the management of giant cell arteritis in routine clinical practice. *Rheumatology* 2018;57:112–9.
22 Hop H, Mulder DJ, Sandovici M, *et al*. Ultrasonography can potentially be the first choice of imaging in suspected extracranial GCA. *Ann Rheum Dis* 2019;78:181–2.
23 Nielsen BD, Hansen IT, Keller KK, *et al*. Ultrasound diagnosis of large vessel inflammation in new-onset treatment-naïve GCA patients using fluorine-18-fluorodeoxyglucose PET/CT as the reference standard—a prospective study of 86 patients suspected of GCA. *Arthritis & Rheumatology* 2018;70.
24 Aschwanden M, Schegk E, Imfeld S, *et al*. Vessel wall plasticity in large vessel giant cell arteritis: an ultrasound follow-up study. *Rheumatology* 2019;58:792–7.

TRANSLATIONAL SCIENCE

Interleukin 1 receptor antagonist (*IL1RN*) gene variants predict radiographic severity of knee osteoarthritis and risk of incident disease

Mukundan Attur ,[1] Hua Zhou,[2] Johathan Samuels,[1] Svetlana Krasnokutsky,[1] Michelle Yau,[3] Jose U Scher,[1] Michael Doherty,[4] Anthony G Wilson ,[5] Jenny Bencardino,[6] Marc Hochberg,[7] Joanne M Jordan,[8] Braxton Mitchell,[7,9] Virginia B Kraus,[10] Steven B Abramson [1]

Handling editor Josef S Smolen

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-216055).

For numbered affiliations see end of article.

Correspondence to
Dr Steven B Abramson, Department of Medicine, Division of Rheumatology, New York University School of Medicine, New York, NY 10016, USA;
Stevenb.Abramson@nyumc.org

MA and HZ are joint first authors.

Received 19 July 2019
Revised 12 November 2019
Accepted 1 December 2019
Published Online First
9 December 2019

To cite: Attur M, Zhou H, Samuels J, *et al*. *Ann Rheum Dis* 2020;**79**:400–407.

ABSTRACT

Objective In these studies, we examined the association of single nucleotide polymorphisms (SNPs) of the *IL1RN* gene with radiographic severity of symptomatic knee osteoarthritis (SKOA) and the risk of incident OA. We also explored these genetic polymorphisms in patients with new onset rheumatoid arthritis (RA).

Methods Over 1000 subjects who met American College of Rheumatology criteria for tibiofemoral OA were selected from three independent, National Institute of Health (NIH)-funded cohorts. CTA and TTG haplotypes formed from three SNPs of the *IL1RN* gene (rs419598, rs315952, rs9005) were assessed for association with radiographic severity, and risk for incident radiographic OA (rOA) in a nested case–control cohort. These *IL1RN* haplotypes were also assessed for association with disease activity (DAS28) and plasma inflammatory markers in patients with RA.

Results Carriage of the *IL1RN* TTG haplotype was associated with increased odds of more severe rOA compared with age-matched, sex-matched and body mass index-matched individuals. Examination of the osteoarthritis initiative *Incidence Subcohort* demonstrated that carriage of the TTG haplotype was associated with 4.1-fold (p=0.001) increased odds of incident rOA. Plasma IL-1Ra levels were lower in TTG carriers, while chondrocytes from TTG carriers exhibited decreased secretion of IL-1Ra. In patients with RA, the TTG haplotype was associated with increased DAS28, decreased plasma IL-1Ra and elevations of plasma inflammatory markers (hsCRP, interleukin 6 (IL-6)).

Conclusion Carriage of the *IL1RN* TTG haplotype is associated with more severe rOA, increased risk for incident OA, and increased evidence of inflammation in RA. These data suggest that the *IL1RN* TTG risk haplotype, associated with decreased IL-1Ra plasma levels, impairs endogenous 'anti-inflammatory' mechanisms.

Key messages

What is already known about this subject?

- Prior genetic studies have not identified any single causal locus with large effects on osteoarthritis (OA), but rather support the polygenic nature of the disease, consistent with the contribution of multiple variants with small effect sizes to variation in OA susceptibility or severity. The *IL-1* gene cluster region has been associated with susceptibility to OA in various joints, but the results have been inconsistent.

What does this study add?

- ➢The *IL1RN* associations that we describe in over 1000 patients with symptomatic knee OA are compelling because the risk haplotype is highly prevalent and has a large, biologically consistent effect on age-dependent radiographic severity or risk of incident disease.
- Our demonstration that the *IL1RN* risk haplotype is associated with more severe rheumatoid arthritis (RA) extends the biological implications to other chronic inflammatory conditions.
- From a pathogenic perspective, the association of the *IL1RN* TTG risk haplotype with decreased plasma IL-1Ra and increased IL-6/hsCRP suggests that carriers of the *IL1RN* TTG haplotype experience more severe and earlier disease due to genetically determined impaired 'anti-inflammatory' mechanisms.

INTRODUCTION

Osteoarthritis (OA) is characterised by focal loss of joint articular cartilage, osteophyte formation and subchondral bone remodelling. The production of interleukin 1 (IL-1β) and other mediators produced by cartilage and synovium induce a state of chronic low-grade inflammation that has been suggested to contribute to disease pathogenesis.[1–4] Multiple genome-wide associations and candidate gene studies have identified genetic variants involved in the pathogenesis of OA,[5–9] including variants in *ALDH1A2*, *GDF5*, *VDR*, *IGF-1*, *COL11A1* and *VEGF*. However, genetic studies have not identified any single causal locus, but rather are consistent with the contribution of multiple variants with small effect sizes to variation in OA susceptibility or severity.[10–12]

We have previously examined 15 single nucleotide polymorphisms (SNPs) in six inflammatory response genes, including those for *IL-1α*, *IL-1β*, *IL-1RN*, *TNFα*, *IL-10*, oestrogen receptor 1 (*ESR1*)

Key messages

How might this impact on clinical practice or future developments?

- Drug development in OA would benefit from genetic biomarkers that identify individuals at greater risk for more severe or incident OA.
- Stratification by *IL1RN* risk haplotype in future clinical trial design and personalised medicine strategies could identify subsets of anti-IL1 responders/non-responders based on *IL1RN* risk haplotypes, as has been described in juvenile systemic arthritis.
- Finally, the understanding of the pathogenic mechanisms of *IL1RN* variants that impair effective endogenous anti-inflammatory mechanisms in OA and RA could lead to the identification of novel targets for treatment.

and determined whether polymorphisms of these genes could predict risk for radiographic knee OA severity. We found that radiographic severity was associated only with a three SNP haplotype (rs419598, rs315952 and rs9005) of *IL1RN*, the product of which is IL-1Ra.[13] The goal of this study was to validate these findings in over 1000 additional individuals with or at risk for knee OA and to determine whether the findings extended to patients with rheumatoid arthritis (RA).

METHODS

Participants with symptomatic knee OA

We assembled 1066 subjects from three independent cohorts of individuals with or at risk for knee OA. Participants met clinical (American College of Rheumatology) and radiographic criteria for tibiofemoral OA (Kellgren-Lawrence (KL) score ≥1); all had body mass index (BMI) <33 kg/m^2 (see online supplementary file 1). Using these eligibility criteria, we established a study population by including 300–400 subjects from each cohort with the goal of reducing phenotypic heterogeneity across populations. Radiographs were scored for tibiofemoral KL grade (0–4) and minimal medial joint space width (mJSW).[1 14–16]

New York University OA cohort

To validate our original observation linking *IL1RN* haplotypes to OA severity from the 'founding' cohort of 80 New York University (NYU) and 50 Duke symptomatic knee osteoarthritis (SKOA) patients,[13 17] we recruited and followed 372 additional SKOA patients between 2008 and 2016. Individuals who comprised the 'founding' cohort are not included in this study (NYUSoM IRB approved no: # i05-131 and i12-03682).

Genetics of Generalized Osteoarthritis

We applied the same inclusion/exclusion criteria to select a subset of participants in the Genetics of Generalized Osteoarthritis (GOGO) study from Duke University,[14] and identified 339 individuals who met the eligibility criteria. None of the GOGO patients selected for this study were among the participants included in the previously reported 'founding' cohort.[13]

Osteoarthritis initiative

We applied identical criteria to select a subject subset from the osteoarthritis initiative (OAI), an observational cohort study focused on identifying genetic and clinical risk factors, imaging and biochemical biomarkers for development and progression of knee OA. The OAI study recruited individuals divided into two subcohorts, 'Progression' and 'Incidence'; inclusion and exclusion criteria for entry into the *Progression and Incidence Subcohorts* are available at http://oai.epi-ucsf.org/datarelease/.

Risk for SKOA

Using the OAI *Incidence Subcohort*, we performed a nested case–control study to assess the risk of incident disease. We identified 101 cases who developed either radiographic or symptomatic tibiofemoral radiographic knee OA within 2–4 years of baseline, and compared 101 controls who did not develop either frequent knee pain or radiographic tibiofemoral OA (>KL1) over a minimum of 4 years and for up to 96 months of follow-up, matched for gender, age and BMI at baseline visit (see online supplementary methods).

NYU new-onset RA cohort

All patients met the American College of Rheumatology/European League Against Rheumatism 2010 classification criteria for RA.[18] Enrolled patients were seropositive: rheumatoid factor (95%); anti-citrullinated protein antibodies (100%). New-onset RA was defined as disease duration of a minimum of 6 weeks and up to 6 months since diagnosis, and absence of any treatment with disease-modifying anti-rheumatic drugs (DMARDs), biological therapy or steroids (ever) as we have described.[19] Plasma samples from 145 RA subjects were selected for analysis. Clinical assessments included tender and swollen 28-joint counts, patient global disease activity assessment (0–100), and ESR to enable calculation of the DAS28-ESR.[20 21]

Haplotype determination

Since all three SNPs (rs419598, rs315952 and rs9005) are in the *IL1RN* gene, we evaluated haplotype effects on radiographic severity as described in our previous publication.[13] All cases and controls were genotyped for the same set of SNP markers (rs419598, rs315952 and rs9005) in the *IL1RN* gene. Of the nine potential haplotypes that could be constructed from these three SNPs, two occurred with a frequency that were >1% (haplotypes CTA and TTG). Both CTA and TTG are found on the same locus. Specifically, 61.7% of subjects could be unambiguously inferred to carry 0, 1, or 2 copies of the TTG haplotype, and 12% of subjects could be unambiguously inferred to carry 1 or 2 copies of the CTA haplotype. Throughout this report, we denote the TTG-0 or TTG-1 or TTG-2 haplotype groups, to represent carriers of 0, 1, or 2 copies of the IL1RN TTG haplotype generated from the 3 IL1RN SNPs (rs419598, rs315952 and rs9005). The linkage disequilibrium parameters D′ and r2 for IL1RN SNPs rs419598, rs315952 and rs9005 are shown in online supplementary table 1 for all three cohorts. In the GOGO cohort, rs9005 was not directly genotyped but was imputed with high quality (INFO >0.8). For consistency, we used the most probable imputed genotypes for all three SNPs to generate IL1RN haplotypes. For both rs315952 and rs419598 SNPs genotype concordance was excellent (r^2=0.981 and 0.976, respectively).

Genetics and molecular analysis

Genotyping, cell culture assays and ELISA were performed as described in online supplementary file 1.

Statistical analyses

Primary analyses evaluated associations between haplotypes and radiographic severity. Genotype associations with radiographic

Table 1 Demographic and radiographic characteristic features of symptomatic knee osteoarthritis patients from NYU, GOGO and OAI cohorts

Cohort	Age in years	Sex (% females)	BMI	Ethnicity (% Caucasians)	KL 3/4 (%)	mJSW in mm	Haplotype % frequencies				
							TTG-0	CTA-1/2 (TTG-0)	TTG-1 (CTA-0)	TTG-2	CTA-1 (TTG-1)
NYU (n=372)	61.4 (±10.4)	62.7	26.41 (±3.52)	64	48.9	3.21 (±1.53)	22.3	8.0	46.0	15.3	10.5
GOGO (n=339)	67.0 (±8.2)	75.5	26.8 (±3.40)	100	29.8	3.35 (±1.4)	32.0	20.0	41.0	27.0	0
OAI (n=355)	61.6 (±9.0)	56.9	29.95 (±4.90)	80	57.7	3.54 (±1.65)	24.2	9.3	40.0	17.2	16.4
Combined (n=1066)	62.7 (±9.9)	64.4	27.36 (±4.15)	68.5	43.2	3.24 (±1.61)	26.3	12.0	42.2	19.5	9.0

Details are shown of the mean (±SD) age, BMI and mJSW in millimetres (mm), as well as percentage of females, Caucasians, radiographic KL score 3 or 4 distribution and *IL1RN* TTG haplotype frequency distribution. The TTG-0 groups also include CTA-1/2 haplotype group. Haplotypes TTG-0 or TTG-1 or TTG-2, respectively, represent carriers of 0 or 1 or 2 copies of *IL1RN* haplotype produced using 3 *IL1RN* single nucleotide polymorphisms (rs419598, rs315952 and rs9005).
BMI, body mass index; GOGO, Genetics of Generalized Osteoarthritis study, Duke University; KL, Kellgren-Lawrence; mJSW, minimal medial joint space width; NYU, New York University School of Medicine; OAI, osteoarthritis initiative.

severity were determined using Fisher's exact test, adjusted using the false discovery rate, where appropriate.

For a continuous trait outcome, mJSW versus age, we used a regression model. Age and mJSW correlation were plotted, and at each age interval, the likelihood of mJSW was calculated with a 95% CI (for additional information see online supplementary file 1).

RESULTS

Frequency of IL1RN haplotypes

The clinical, genetic and demographic parameters in the three cohorts are shown in table 1. The frequencies of the *IL1RN* TTG haplotypes, based on SNPs rs419598, rs315952 and rs9005, were similar across the three cohorts. For the combined cohort of 1066 participants, the frequencies of TTG-0, TTG-1 and TTG-2 were 26.3%, 42.2% and 19.5%, respectively. The overall frequency of the CTA-1 or CTA-2 haplotype was 12.0% but varied across cohorts (NYU 7%; GOGO 20%; OAI 9%). Approximately 30% of the TTG-0 haplotype subjects in the combined cohort were CTA carriers.

IL1RN TTG haplotype is associated radiographic severity

We first examined *IL1RN* TTG-0 versus TTG haplotypes (TTG-1 and TTG-2) for association with radiographic OA (rOA) severity as reported by KL scores and mean minimal medial radiographic joint space width (mJSW). As shown in a Forest plot (figure 1A), the *IL1RN* TTG haplotype was associated with an increased odds of more severe (KL 3/4 vs KL 1/2) radiographic knee OA compared with age-matched, sex-matched and BMI-matched knee OA patients with TTG-0 (OR of 1.83; 95% CI 1.36 to 2.46; p=0.0003). In the GOGO cohort, the TTG haplotype was associated with increased odds of KL 3/4 OA, which did not achieve significance. Of note, there was a lower percentage of participants with KL 3/4 OA severity (29.8%) in GOGO

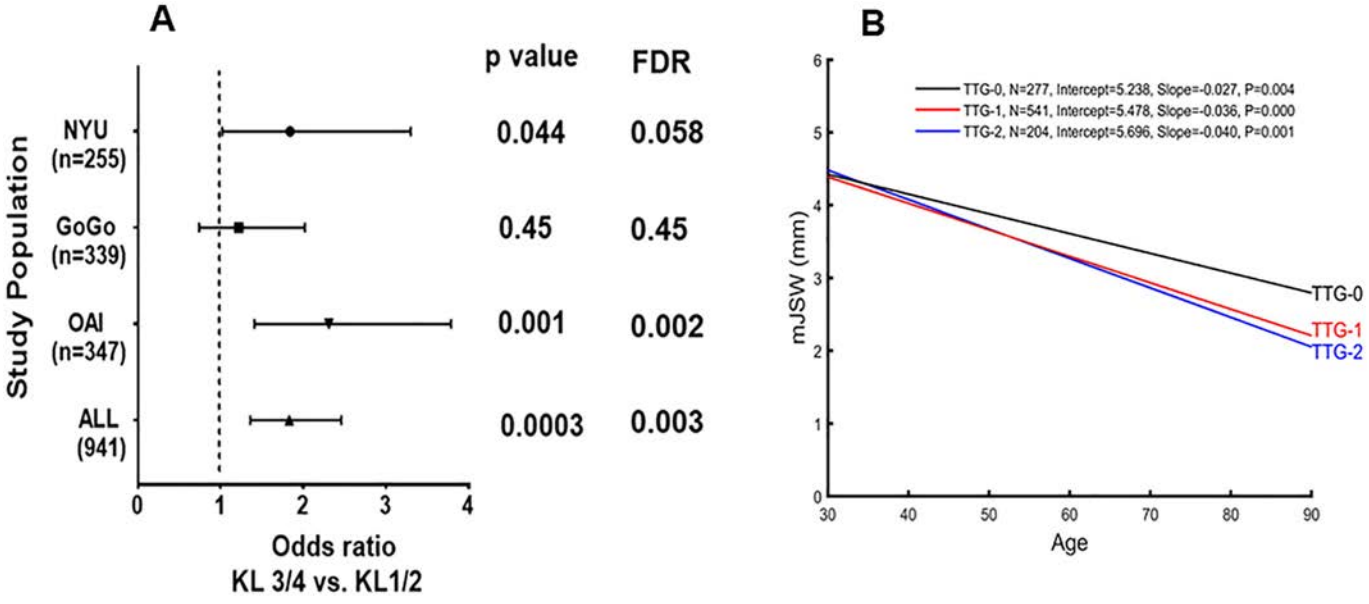


Figure 1 Association of *IL1RN* TTG haplotypes with radiographic severity. (A) Forest plot displaying association of *IL-1RN* haplotypes (TTG-0 vs TTG-1/2) with radiographic severity in symptomatic knee osteoarthritis (OA) patients in three cohorts. Study-specific estimates of ORs with 95% CIs between severity of knee OA defined as Kellgren-Lawrence (KL) 1/2 versus KL 3/4 for haplotype rs419598, rs315952 and rs9005 'T-T-G' are shown for three independent and all three combined cohort. (B) Association of radiographic minimal medial joint space width (mJSW), age and *IL1RN* haplotypes in NYU, OAI and GOGO cohorts. Influence of IL-1 receptor antagonist (IL1RN) haplotypes on the age relationship to mJSW of knee OA. Carriers of either TTG-1 or TTG-2 compared with TTG-0 had narrower JSW (mm) at each age (years) studied. The joint space width (JSW) of each knee in patients with knee OA who do not (TTG-0) or do carry the *IL1RN* TTG haplotype is plotted relative to age, and the regression line is shown for JSW relative to age. The figure shows the linear regression line for each of the *IL1RN* risk haplotypes. OA patients 982 out of 1066 from three cohorts are represented. GOGO, Genetics of Generalized Osteoarthritis; FDR, false discovery rate; NYU, New York University; OAI, osteoarthritis initiative.

Table 2 Association of IL1RN haplotype (TTG) with radiographic mean minimal medial joint space width (mJSW) in three combined (NYU, GOGO and OAI) cohorts of symptomatic knee osteoarthritis patients

Cohorts	TTG-0	TTG-1	TTG-2	Beta (95% CI)	P value	FDR
NYU (n=372)	3.43 (±1.44) (n=83)	3.28 (±1.45) (n=209)	2.60 (±1.70) (n=57)	−0.39 (−0.64 to −0.13)	0.0030	0.0046
GOGO (n=339)	3.67 (±1.31) (n=111)	3.08 (±1.51) (n=138)	3.35 (±1.38) (n=90)	−0.18 (−0.38 to 0.02)	0.0049	0.0950
OAI (n=355)	3.40 (±1.46) (n=86)	3.20 (±1.63) (n=200)	3.09 (±1.81) (n=61)	−0.16 (−0.42 to 0.10)	0.4787	0.2364
All (n=1066)	3.52 (±1.40) (n=280)	3.20 (±1.53) (n=547)	3.07 (±1.63) (n=208)	−0.23 (−0.37 to −0.10)	0.0023	0.0021
All meta-analysis (n=1066)	–	–	–	−0.23 (−0.37 to −0.10)	0.0008	0.0021

mJSW data are presented in millimetres as mean (±SD) unless otherwise indicated; number of subjects in each group is represented in square brackets below each value.
Haplotypes TTG-0 or TTG-1 or TTG-2, respectively, represent carriers of 0 or 1 or 2 copies of *IL1RN* haplotype produced using 3 *IL1RN* single nucleotide polymorphisms (rs419598, rs315952 and rs9005). Linear regression model (beta and 95% CI) was performed and p value was adjusted by FDR. The last row indicates the meta-analysis of all three cohorts by including cohort as a covariate.
FDR, false discovery rate; GOGO, Genetics of Generalized Osteoarthritis; NYU, New York University; OAI, osteoarthritis initiative.

compared with the NYU (48.9%) and OAI (57.7%) cohorts, which may have reduced the statistical power of the test. We also note the higher frequency of the 'protective' CTA haplotype in the GOGO population (GOGO 20%; NYU 7%; OAI 9%), which could account for the smaller percentage of subjects with severe rOA, as we have reported.[8]

We next assessed radiographic severity by minimal medial joint space width (mJSW). Table 2 shows that relative to TTG-0, carriage of TTG-1 or TTG-2 was associated with a TTG 'dose-dependent' decrease in mJSW in each individual population. Linear regression analysis confirmed that compared with TTG-0, the TTG-1 or TTG-2 haplotypes were significantly associated with decreased mJSW (online supplementary table 2A). This TTG dose effect was also observed for KL severity (online supplementary table 2B). The risk haplotype TTG carriers did not associate with either Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) or Visual Analog Scale (VAS) pain in the NYU or OAI cohort.

We next performed a more detailed analysis of the effects of different combinations of CTA and TTG haplotypes on mJSW. As shown in table 3, any combination of CTA-1 or CTA-2 was associated with wider mJSW compared with TTG-2 or TTG-1. For example, the mJSW for CTA-2 and CTA-1 or CTA-2 carriers was 3.67 (1.3) and 3.51 (1.31), respectively. In comparison, the mean mJSW for TTG-2 and TTG-1 or TTG-2 carriers was 3.07 (1.62) and 3.19 (1.57), respectively. The differences between CTA and TTG mJSW were significant after adjustment for common covariates age, sex and BMI (table 3).

IL1RN TTG haplotype predicts age-related rOA

We next evaluated the interaction between age, mJSW and *IL1RN* genotypes relative to radiographic severity. As shown in figure 1B, linear regression analysis demonstrated that carriers of either TTG-1 or TTG-2 compared with TTG-0 had narrower JSW (mm) at each age studied. At age 70, for example, mean mJSW was 3.34 mm in TTG-0 versus 2.86 mm in TTG-2 ($p<0.005$).

We analysed whether rOA was associated with *IL1RN* risk haplotype after adjustment for risk covariates (age, sex and BMI) in the regression model. As shown in online supplementary table 3, IL1RN risk haplotype carriers (TTG-1 or 2) had significantly narrower tibiofemoral mJSW compared with TTG-0 carriers; the association remained significant after adjustment for the covariates. In gender-specific analyses, we show that both male and female carriers of either TTG-1 or TTG-2 carriers had narrower mJSW compared with TTG-0 (online supplementary table 4). In addition, among Blacks and Hispanics TTG-1 or TTG-2 carriers had narrower mJSW compared with TTG-0 (online supplementary tables 5 and 6) with a p value of 0.05–0.10.

IL1RN TTG haplotype predicts the risk of incident rOA

We examined participants from the *Incidence Subcohort* of the OAI, selecting the subgroup without clinical or radiographic evidence of knee OA at baseline. We identified 101 cases who developed either radiographic or symptomatic tibiofemoral radiographic knee OA within 2–4 years of baseline. Using a nested case–control approach, we selected 101 controls from the OAI *Incidence Subcohort* who did not develop either pain or radiographic tibiofemoral OA (>KL1) over a similar period matched for age, sex and BMI. These subjects were followed for 2–8 years. Table 4 shows that the presence of the *IL1RN* TTG-2 haplotype significantly increased the risk of incident knee rOA (OR=4.13 (1.75–9.72); p=0.001). After adjustment for age, sex and BMI with a logistic regression model, carriage of the TTG haplotype remained positively and significantly associated with incident rOA (beta coefficient=1.38; 95% CI 0.48 to 2.28; p=0.002).

TTG risk haplotype is associated with decreased plasma IL-1Ra levels in patients with OA

Genetic variants of *IL1RN* have been associated with plasma levels of IL-1Ra and may regulate intracellular IL-1Ra protein trafficking.[22–25] In our studies, mean plasma IL1Ra protein concentrations in TTG-2 carriers were lower than in CTA-2 carriers (346.50 vs 479.45 pg/mL, p=0.05, respectively). In this subset of patients, carriers of the CTA-2 haplotype had wider mean (SD) mJSW than did TTG-2 carriers (3.28 (1.46) vs 2.60 (1.67) mm, p=0.046, age-adjusted, sex-adjusted and BMI-adjusted). This was despite the fact that CTA-2 carriers were a mean 6 years *older* than the TTG-2 carriers (68.94 (9.92) vs 62.33 (10.96) years, p=0.08), consistent with an age-evident 'protective' effect of CTA on rOA.

We also performed a causal analysis to determine relationships among *IL1RN* TTG, CTA haplotypes, age, sex, BMI, IL-1Ra and mJSW. As shown in online supplementary figure 2, the CTA

Table 3 *IL1RN* TTG-1/2 carriers consistently associated with narrower joint space width in combined cohorts

	CTA-2	TTG-2	P value	P value_ ASB_ Adj	CTA-1/2 (TTG-0)	TTG-2	P value	P value_ ASB_ Adj	CTA-1/2 (TTG-0)	TTG-1/2 (CTA-0)	P value	P value_ ASB_ Adj
mJSW, mm	3.67 (±1.30)	3.07 (±1.62)	**0.0063**	**0.0072**	3.51 (±1.31)	3.07 (±1.62)	**0.0106**	**0.0124**	3.51 (±1.31)	3.19 (±1.57)	**0.0304**	**0.0123**
Age, years	65.01 (±8.24)	65.03 (±9.29)	0.9881	–	64.97 (±8.95)	65.03 (±9.29)	0.9498	–	64.97 (±8.95)	63.64 (±9.85)	0.1571	–
BMI	27.15 (±4.02)	27.74 (±3.98)	0.2933	–	27.31 (±3.66)	27.74 (±3.98)	0.3272	–	27.31 (±3.66)	27.55 (±4.20)	0.5478	–
Sex (% male)	36.76%	35.1%	–	–	33.6%	35.1%	–	–	33.6%	33.4%	–	–

Details are shown of the mean (±SD) age, BMI and mJSW in millimetres (mm), as well as percentage of males in each haplotype group in combined cohort (including New York University, Genetics of Generalized Osteoarthritis and osteoarthritis initiative). Haplotypes TTG-0 or TTG-1 or TTG-2, respectively, represent carriers of 0 or 1 or 2 copies of *IL1RN* haplotype produced using 3 *IL1RN* single nucleotide polymorphisms (rs419598, rs315952 and rs9005). Data are presented as mean (±SD) unless otherwise indicated. P values were determined by two-sample t-test and adjusted for ASB.
Bold indicates significant p values.
ASB, age, sex and BMI; BMI, body mass index; mJSW, minimal joint space width.

haplotype and BMI, but not age, are independently associated with plasma IL-1Ra. The causal analysis also indicated that the TTG haplotype directly associated with mJSW. As expected, both age and BMI associated with mJSW, and these effects were independent of the TTG haplotype.

IL1RN haplotypes in patients with RA

We examined plasma samples from new-onset, DMARD-untreated patients with RA, followed at NYU.[20 21] As shown in figure 2, carriers of the TTG risk haplotype exhibited lower levels of plasma IL-1Ra and the soluble IL-6 receptor antagonist alpha (sIL-6Rα) than age-matched, BMI-matched and sex-matched individuals with RA. Conversely, in TTG carriers plasma IL-6 and hsCRP were higher. Clinically, carriers of the TTG haplotype exhibited greater disease activity (DAS28).

IL1RN haplotypes affect chondrocytes production of IL-1Ra

We next examined the relationship between TTG haplotypes and IL-1Ra production by chondrocytes. Chondrocytes were isolated from patients undergoing total joint replacement surgery at NYU, as described.[26] Cell lysates and matched supernatants were analysed for IL-1Ra protein concentrations after 24 hours culture in the presence or absence of IL-1β. As shown in online supplementary table 7, following exposure to IL-β, basal levels of secreted IL-1Ra did not increase in TTG carriers, whereas *intracellular* concentrations of IL-1Ra in TTG chondrocytes were markedly increased. In contrast, chondrocytes obtained from TTG-0 individuals significantly increased the production of both intracellular and extracellular IL-1Ra following stimulation with IL-1β.

DISCUSSION

The *IL-1* gene cluster region has been associated with susceptibility to OA in various joints, but the results have been inconsistent.[27–32] In this study of more than 1000 individuals with SKOA, we show that carriers of the *IL1RN* CTA haplotype (rs419598, rs315952 and rs9005) exhibit decreased age-dependent radiographic severity. Conversely, the TTG haplotype is associated with more severe rOA. Moreover, we demonstrate that the *IL1RN* TTG haplotype significantly increased the risk for incident tibiofemoral knee OA.

These results are consistent with our previous report that CTA in a large meta-analysis associated with less severe radiographic severity of knee OA.[17] We note that in the genome-wide association study of OA using the UK Biobank, individual associations of each *IL1RN* SNP did not associate with knee OA at the genome-wide threshold ($p<5\times10^{-8}$).[33] Similarly, in our study, *IL1RN* individual SNPs did not associate with knee OA (either KL or JSW).[17] However, only the *IL1RN* haplotype, not tested in the UK Biobank study, was associated with more severe rOA in our studies. Another difference is the use of patient and/or hospital reported OA knee cases in the UK Biobank study, rather than radiographically confirmed SKOA as in our cohort, that may have resulted in a more heterogenous population of OA cases in the UK cohort.[6 33]

We also tested for association between the *IL1RN* haplotypes and radiographic progression, but neither of these associations were statistically significant. This is in contrast to the studies by Wu *et al*, who reported that the *IL1RN* TTG haplotype associated with change in KL over 4–11 years).[34] Therefore, the lack of an association with progression in our studies could be a consequence of low power or insufficient years of follow-up. Alternatively, it is possible that exposures (eg, genotype) that

Table 4 *IL1RN* TTG haplotype increases risk of incident osteoarthritis (OA)

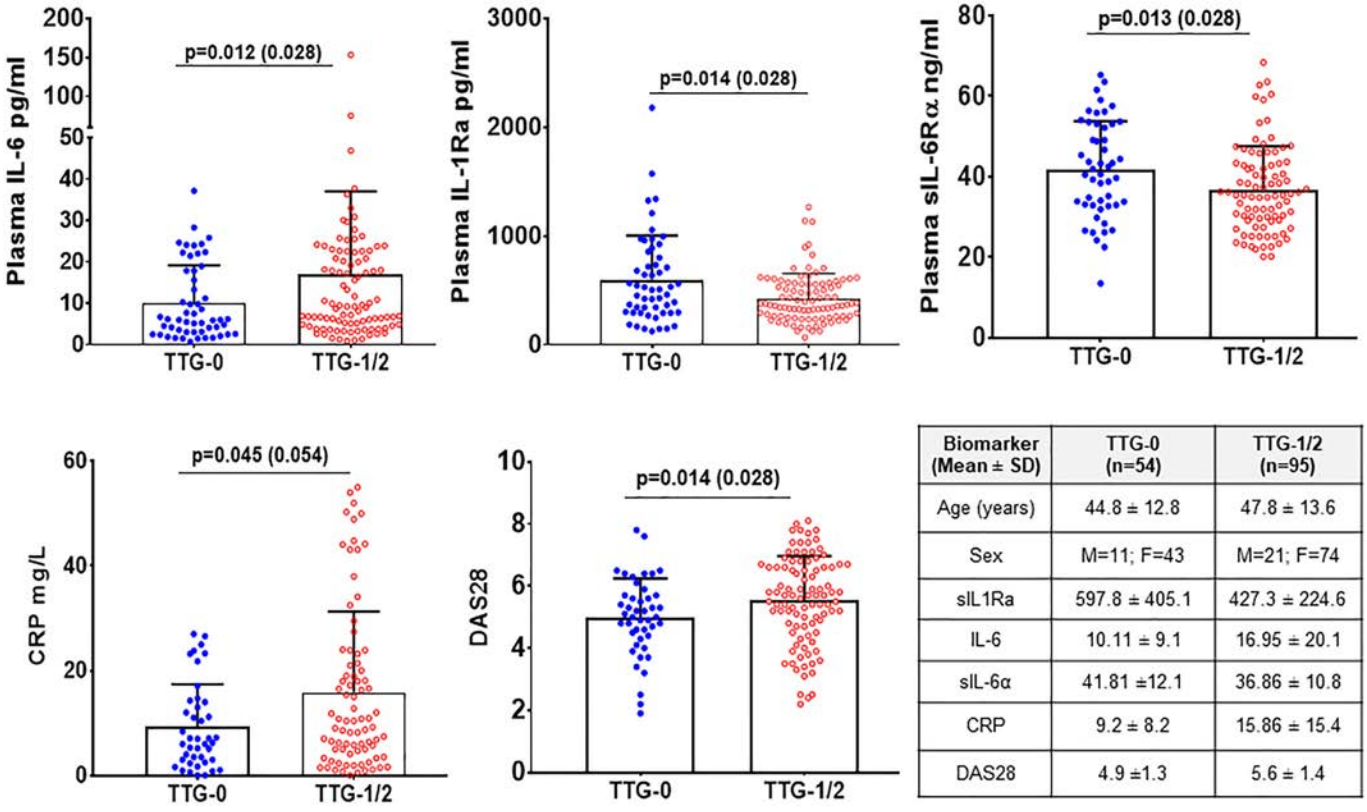
	Age (years)	BMI	Sex	TTG-2	TTG-0	OR (95% CI); P value	Beta ASB adjusted
Cases	62.6±8.9	26.4±3.3	M=31; F=70	48 (M=14; F=34)	16 (M=7; F=9)	4.13 (1.75–9.72); **0.001**	1.38 (0.48–2.28); **0.002**
Controls	62.6±8.8	26.3±3.3	M=31; F=70	16 (M=6; F=10)	22 (M=0; F=22)		

Development of incident OA in cases was defined as development of frequent knee pain and radiographic OA (KL ≥1 or 2) in the same knee or in bilateral knees. Controls were individuals whose baseline Kellgren-Lawrence (KL)=0 or 1 did not change at follow-up AND who did not develop frequent pain in either knee at 24, 36, 48, 72 and 96 months. Cases and controls were matched for age, sex and BMI. Estimates of OR with 95% CIs between severity of knee OA defined as KL 1/2 versus KL 3/4 for haplotype rs419598, rs315952 and rs9005 'T-T-G' are shown. Haplotypes TTG-0 or TTG-1 or TTG-2, respectively, represent carriers of 0 or 1 or 2 copies of *IL1RN* haplotype produced using 3 *IL1RN* SNPs (rs419598, rs315952 and rs9005). Data for age and BMI are presented as mean±SD; data for sex are presented as N of male and female. The ORs of patients falling into case or control groups versus *IL1RN* haplotype were calculated using Fisher's exact test. Beta coefficient (and 95% CI) from logistic regressions were adjusted for ASB.
ASB, age, sex and BMI; BMI, body mass index.

increase disease susceptibility may also promote progression but such an association could be hard to detect because both progressors and non-progressors may already be enriched for the susceptibility genotype. This is a form of selection bias known as 'collider bias'.

The association of *IL1RN* haplotypes with increased rOA at earlier age and the risk of incident disease may have clinical implications. Drug development in OA would benefit from genetic biomarkers that identify individuals at greater risk for more severe or incident OA.[35] Stratification by *IL1RN* risk haplotype in future clinical trial design could identify subsets of anti-IL1 responders/non-responders based on *IL1RN* risk haplotypes, as has been described in juvenile systemic arthritis.[36]

What might be the biological explanation for the '*yin*/*yang*' genetic effects of CTA versus TTG on rOA? We have previously shown that individuals carrying the *IL1RN* CTA (TTG-0) haplotype had significantly lower synovial fluid levels of IL-10 and showed a trend towards lower levels of IL-1β and IL-6.[13] We here report that in patients with both OA and RA, carriers of the TTG haplotype exhibit reduced plasma levels of IL-1Ra compared with CTA carriers. We provide evidence in chondrocytes that this may result from decreased secretion of IL-1Ra protein. Although our studies focused on cartilage, the source of IL-1Ra in the synovial joint fluid could be from various tissues in the joint, including inflamed synovium. We postulate that the greater severity of rOA in carriers of the TTG haplotype results



Biomarker (Mean ± SD)	TTG-0 (n=54)	TTG-1/2 (n=95)
Age (years)	44.8 ± 12.8	47.8 ± 13.6
Sex	M=11; F=43	M=21; F=74
sIL1Ra	597.8 ± 405.1	427.3 ± 224.6
IL-6	10.11 ± 9.1	16.95 ± 20.1
sIL-6α	41.81 ±12.1	36.86 ± 10.8
CRP	9.2 ± 8.2	15.86 ± 15.4
DAS28	4.9 ±1.3	5.6 ± 1.4

Figure 2 IL1RN TTG-risk haplotype carriers have decreased IL-1Ra, soluble IL-6Rα and increased IL-6, CRP accompanied by increased disease activity (DAS28) in rheumatoid arthritis patients. Plasma levels of biomarkers were determined using ELISA as described in methods. The mean (SD) age, sex and biomarkers and number of subjects in each haplotype group are shown in the table. Each dot represents individual sample. The solid horizontal bar in each group represents the mean and the vertical bar represents the positive SD values. Mann-Whitney U test was used to analyse significance difference between specific haplotype groups (TTG-1/2) with haplotype of TTG-0 groups. The p value and false discovery rate adjusted values (in brackets) are shown in the figures. CRP, C-reactive protein.

from impaired antagonism of chronic inflammatory IL-1β-driven processes.[1 2]

In these studies, we also asked whether the association of the TTG haplotype with more severe disease was limited to OA, or could be demonstrated in patients with new onset RA. We found that in RA, as in OA, plasma levels of IL-1Ra were decreased in TTG carriers, and this was accompanied by *increased* plasma IL-6 and hsCRP in association with *increased* clinical disease activity (DAS28).

With regard to limitations, our studies were restricted to SKOA, since standardised radiographs of other joints were not available in each of the study cohorts. Therefore, the risk conferred by the *IL1RN* risk haplotypes to individuals with OA of the hip, hands, and/or the spine will need assessment in future studies. In addition, we note that participants enrolled in our three cohorts were predominantly North American Caucasian. However, although the numbers were small, subset analysis of Black and Hispanic subjects indicated a trend towards increased rOA severity in each subset (online supplementary tables 5 and 6). Thus, confirmation of these findings in Black, Asian and Hispanic populations will require future studies.

CONCLUSION

In summary, we demonstrate that the *IL1RN* TTG haplotype identifies a subset of individuals with knee OA who are at increased risk for age-dependent rOA and increased risk for incident OA. Evidence for increased serological and clinical markers of disease activity in TTG carriers is also provided in new onset RA. We postulate that carriers of the *IL1RN* TTG haplotype experience more severe disease due to genetically determined impaired 'anti-inflammatory' mechanisms.

Author affiliations
[1]Department of Medicine, Division of Rheumatology, New York University School of Medicine, New York, New York, USA
[2]Applied Bioinformatics Laboratories, New York University School of Medicine, New York, New York, USA
[3]Department of Epidemiology and Public Health, University of Maryland Baltimore, Baltimore, Maryland, USA
[4]Academic Rheumatology, University of Nottingham, Nottingham, UK
[5]University College Dublin, UCD School of Medicine and Medical Science, Conway Institute, Dublin, Ireland
[6]Radiology, New York University School of Medicine, New York, New York, USA
[7]Medicine/Epidemiology and Public Health, University of North Carolina, Chaple Hill, North Carolina, USA
[8]Departments of Medicine and Epidemiology, Thurston Arthritis Research Center, Chapel Hill, North Carolina, USA
[9]Geriatrics Research and Education Clinical Center, Baltimore VA Medical Center, Baltimore, Maryland, USA
[10]Medicine, Duke University Medical Center, Durham, North Carolina, USA

Acknowledgements The NYU and OAI cohort bio-specimens were collected as part of a grant to SBA from the US National Institutes of Health (R01AR052873). Ascertainment of GOGO study participants was made possible by funding from GlaxoSmithKline. The OAI is a public–private partnership comprised five contracts (N01-AR-2-2258; N01-AR-2-2259; N01-AR-2-2260; N01-AR-2-2261; N01-AR-2-2262) funded by the National Institutes of Health, a branch of the US Department of Health and Human Services, and conducted by the OAI Study Investigators. Private funding partners include Merck Research Laboratories; Novartis Pharmaceuticals Corporation, GlaxoSmithKline; and Pfizer. Private sector funding for the OAI was managed by the Foundation for the National Institutes of Health. This manuscript was prepared using an OAI public use data set and does not necessarily reflect the opinions or views of the OAI investigators, the NIH, or the private funding partners. This work was also supported by the generous contributions of the William and Lynda Steere Foundation. The authors would like to thank the staff and participants of the NYU Osteoarthritis Biomarkers Study.

Contributors MA, JS, SK, JUS, MH, JMJ, VBK and SBA conceived and designed the study. MA, JS, SK, JUS and JB acquired the data. All authors were involved in analysis and interpretation of the data. All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted. MA, HZ, MY and SBA had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Funding The study was funded in part by a grant to Dr Abramson from the US National Institutes of Health (R01AR052873), by funding from GlaxoSmithKline to support inclusion of GOGO study participants, and by generous contributions from the William and Lynda Steere Foundation. None of the funding sources had any involvement in the design, execution and reporting of this study.

Competing interests MA and SBA have one provisional patent application and another approved patent for the use of inflammatory and genetic biomarkers in predicting at-risk knee OA patients.

Patient consent for publication Not required.

Ethics approval The current study was performed in accordance with the ethical standards of the Declaration of Helsinki 1975, as revised in 2000, and studies were approved by the Institutional Review Board (IRB) of NYU School of Medicine.

Provenance and peer review Not commissioned; externally peer-reviewed.

Data availability statement Patient deidentified- NYU data are freely available upon request. OAI data are available from OAI.epi-uscf.org

ORCID iDs
Mukundan Attur https://orcid.org/0000-0001-7080-8118
Anthony G Wilson http://orcid.org/0000-0003-4855-3926
Steven B Abramson http://orcid.org/0000-0002-0668-6344

REFERENCES

1 Attur M, Krasnokutsky S, Statnikov A, *et al*. Low-Grade inflammation in symptomatic knee osteoarthritis: prognostic value of inflammatory plasma lipids and peripheral blood leukocyte biomarkers. *Arthritis Rheumatol* 2015;67:2905–15.
2 Robinson WH, Lepus CM, Wang Q, *et al*. Low-Grade inflammation as a key mediator of the pathogenesis of osteoarthritis. *Nat Rev Rheumatol* 2016;12:580–92.
3 Watson Levings RS, Broome TA, Smith AD, *et al*. Gene therapy for osteoarthritis: pharmacokinetics of intra-articular self-complementary adeno-associated virus interleukin-1 receptor antagonist delivery in an equine model. *Hum Gene Ther Clin Dev* 2018;29:90–100.
4 Attur MG, Dave M, Cipolletta C, *et al*. Reversal of autocrine and paracrine effects of interleukin 1 (IL-1) in human arthritis by type II IL-1 decoy receptor. potential for pharmacological intervention. *J Biol Chem* 2000;275:40307–15.
5 Gonzalez A, Valdes AM. Big data boost for osteoarthritis genetics. *Nat Rev Rheumatol* 2018;14:387–8.
6 Zengini E, Hatzikotoulas K, Tachmazidou I, *et al*. Genome-Wide analyses using UK Biobank data provide insights into the genetic architecture of osteoarthritis. *Nat Genet* 2018;50:549–58.
7 Rodriguez-Fontenla C, Calaza M, Evangelou E, *et al*. Assessment of osteoarthritis candidate genes in a meta-analysis of nine genome-wide association studies. *Arthritis Rheumatol* 2014;66:940–9.
8 Yau MS, Yerges-Armstrong LM, Liu Y, *et al*. Genome-Wide association study of radiographic knee osteoarthritis in North American Caucasians. *Arthritis Rheumatol* 2017;69:343–51.
9 Styrkarsdottir U, Thorleifsson G, Helgadottir HT, *et al*. Severe osteoarthritis of the hand associates with common variants within the aldh1a2 gene and with rare variants at 1p31. *Nat Genet* 2014;46:498–502.
10 Fernandez-Moreno M, Rego I, Carreira-Garcia V, *et al*. Genetics in osteoarthritis. *Curr Genomics* 2008;9:542–7.
11 Loughlin J. Genetic contribution to osteoarthritis development: current state of evidence. *Curr Opin Rheumatol* 2015;27:284–8.
12 Reynard LN, Loughlin J. The genetics and functional analysis of primary osteoarthritis susceptibility. *Expert Rev Mol Med* 2013;15:e2.
13 Attur M, Wang H-Y, Kraus VB, *et al*. Radiographic severity of knee osteoarthritis is conditional on interleukin 1 receptor antagonist gene variations. *Ann Rheum Dis* 2010;69:856–61.
14 Kraus VB, Jordan JM, Doherty M, *et al*. The genetics of generalized osteoarthritis (GOGO) study: study design and evaluation of osteoarthritis phenotypes. *Osteoarthritis Cartilage* 2007;15:120–7.
15 Mercier C, Piperno M, Vignon E, *et al*. In normal knees, joint space width (JSW) is correlated with the intermargin distance (IMD), a measure of medial tibial plateau alignment. variations in IMD explain variability in JSW in serial radiographs. *Joint Bone Spine* 2013;80:183–7.

16 Attur M, Statnikov A, Samuels J, *et al*. Plasma levels of interleukin-1 receptor antagonist (IL1Ra) predict radiographic progression of symptomatic knee osteoarthritis. *Osteoarthritis Cartilage* 2015;23:1915–24.
17 Kerkhof HJM, Doherty M, Arden NK, *et al*. Large-Scale meta-analysis of interleukin-1 beta and interleukin-1 receptor antagonist polymorphisms on risk of radiographic hip and knee osteoarthritis and severity of knee osteoarthritis. *Osteoarthritis Cartilage* 2011;19:265–71.
18 Aletaha D, Neogi T, Silman AJ, *et al*. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League against rheumatism collaborative initiative. *Arthritis Rheum* 2010;62:2569–81.
19 Scher JU, Sczesnak A, Longman RS, *et al*. Expansion of intestinal Prevotella copri correlates with enhanced susceptibility to arthritis. *eLife* 2013;2:e01202.
20 Jain M, Attur M, Furer V, *et al*. Increased plasma IL-17F levels in rheumatoid arthritis patients are responsive to methotrexate, anti-TNF, and T cell costimulatory modulation. *Inflammation* 2015;38:180–6.
21 Scher JU, Ubeda C, Equinda M, *et al*. Periodontal disease and the oral microbiota in new-onset rheumatoid arthritis. *Arthritis Rheum* 2012;64:3083–94.
22 Rafiq S, Stevens K, Hurst AJ, *et al*. Common genetic variation in the gene encoding interleukin-1-receptor antagonist (IL-1ra) is associated with altered circulating IL-1ra levels. *Genes Immun* 2007;8:344–51.
23 Matteini AM, Li J, Lange EM, *et al*. Novel gene variants predict serum levels of the cytokines IL-18 and IL-1RA in older adults. *Cytokine* 2014;65:10–16.
24 Ramírez-Pérez S, De la Cruz-Mosso U, Hernández-Bello J, *et al*. High expression of interleukine-1 receptor antagonist in rheumatoid arthritis: Association with *IL1RN*2/2* genotype. *Autoimmunity* 2017;50:468–75.
25 Redlitz KH, Yamshchikov VF, Cominelli F. Differential contribution of IL-1ra isoforms to allele-specific IL-1ra mRNA accumulation. *J Interferon Cytokine Res* 2004;24:253–60.
26 Attur M, Yang Q, Shimada K, *et al*. Elevated expression of periostin in human osteoarthritic cartilage and its potential role in matrix degradation *via* matrix metalloproteinase-13. *Faseb J* 2015;29:4107–21.
27 Loughlin J, Dowling B, Mustafa Z, *et al*. Association of the interleukin-1 gene cluster on chromosome 2q13 with knee osteoarthritis. *Arthritis Rheum* 2002;46:1519–27.
28 Moos V, Rudwaleit M, Herzog V, *et al*. Association of genotypes affecting the expression of interleukin-1beta or interleukin-1 receptor antagonist with osteoarthritis. *Arthritis Rheum* 2000;43:2417–22.
29 Smith AJP, Keen LJ, Billingham MJ, *et al*. Extended haplotypes and linkage disequilibrium in the IL1R1-IL1A-IL1B-IL1RN gene cluster: association with knee osteoarthritis. *Genes Immun* 2004;5:451–60.
30 Kanoh T, Hasegawa Y, Masui T, *et al*. Interleukin-1Beta gene polymorphism associated with radiographic signs of osteoarthritis of the knee. *J Orthop Sci* 2008;13:97–100.
31 Moxley G, Han J, Stern AG, *et al*. Potential influence of IL1B haplotype and IL1A-IL1B-IL1RN extended haplotype on hand osteoarthritis risk. *Osteoarthritis Cartilage* 2007;15:1106–12.
32 Moxley G, Meulenbelt I, Chapman K, *et al*. Interleukin-1 region meta-analysis with osteoarthritis phenotypes. *Osteoarthritis Cartilage* 2010;18:200–7.
33 Tachmazidou I, Hatzikotoulas K, Southam L, *et al*. Identification of new therapeutic targets for osteoarthritis through genome-wide analyses of UK Biobank data. *Nat Genet* 2019;51:230–6.
34 Wu X, Kondragunta V, Kornman KS, *et al*. Il-1 receptor antagonist gene as a predictive biomarker of progression of knee osteoarthritis in a population cohort. *Osteoarthritis Cartilage* 2013;21:930–8.
35 Karsdal MA, Christiansen C, Ladel C, *et al*. Osteoarthritis--a case for personalized health care? *Osteoarthritis Cartilage* 2014;22:7–16.
36 Arthur VL, Shuldiner E, Remmers EF, *et al*. *IL1RN* Variation influences both disease susceptibility and response to recombinant human interleukin-1 receptor antagonist therapy in systemic juvenile idiopathic arthritis. *Arthritis Rheumatol* 2018;70:1319–30.

TRANSLATIONAL SCIENCE

Single-cell RNA-seq analysis identifies meniscus progenitors and reveals the progression of meniscus degeneration

Hao Sun ,[1,2] Xingzhao Wen,[2] Hongyi Li,[2] Peihui Wu,[2] Minghui Gu,[2] Xiaoyi Zhao,[2] Ziji Zhang,[2] Shu Hu,[2] Guping Mao,[2] Ruofan Ma,[1] Weiming Liao,[2] Zhiqi Zhang [2]

Handling editor Josef S Smolen

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-215926).

[1]Department of Orthopedics, Sun Yat-Sen Memorial Hospital, Guangzhou, China
[2]Department of Joint Surgery, Sun Yat-sen University First Affiliated Hospital, Guangzhou, China

Correspondence to
Dr Zhiqi Zhang, Department of Joint Surgery, Sun Yat-sen University First Affiliated Hospital, Guangzhou 510080, China; zhzhiqi@mail.sysu.edu.cn
Dr Weiming Liao; liaowmsysu@163.com

HS and XW contributed equally.

Received 28 June 2019
Revised 21 November 2019
Accepted 7 December 2019
Published Online First 23 December 2019

To cite: Sun H, Wen X, Li H, *et al*. *Ann Rheum Dis* 2020;**79**:408–417.

ABSTRACT

Objectives The heterogeneity of meniscus cells and the mechanism of meniscus degeneration is not well understood. Here, single-cell RNA sequencing (scRNA-seq) was used to identify various meniscus cell subsets and investigate the mechanism of meniscus degeneration.

Methods scRNA-seq was used to identify cell subsets and their gene signatures in healthy human and degenerated meniscus cells to determine their differentiation relationships and characterise the diversity within specific cell types. Colony-forming, multi-differentiation assays and a mice meniscus injury model were used to identify meniscus progenitor cells. We investigated the role of degenerated meniscus progenitor (DegP) cell clusters during meniscus degeneration using computational analysis and experimental verification.

Results We identified seven clusters in healthy human meniscus, including five empirically defined populations and two novel populations. Pseudotime analysis showed endothelial cells and fibrochondrocyte progenitors (FCP) existed at the pseudospace trajectory start. Melanoma cell adhesion molecule ((MCAM)/CD146) was highly expressed in two clusters. CD146+ meniscus cells differentiated into osteoblasts and adipocytes and formed colonies. We identified changes in the proportions of degenerated meniscus cell clusters and found a cluster specific to degenerative meniscus with progenitor cell characteristics. The reconstruction of four progenitor cell clusters indicated that FCP differentiation into DegP was an aberrant process. Interleukin 1β stimulation in healthy human meniscus cells increased CD318+ cells, while TGFβ1 attenuated the increase in CD318+ cells in degenerated meniscus cells.

Conclusions The identification of meniscus progenitor cells provided new insights into cell-based meniscus tissue engineering, demonstrating a novel mechanism of meniscus degeneration, which contributes to the development of a novel therapeutic strategy.

Key messages

What is already known about this subject?

- The cell types of meniscus contain chondrocyte-like morphology cells and fibroblast-like cells. However, the variety of cell types and corresponding biological markers, as well as the biological targets for the treatment for meniscus degeneration remain elusive.

What does this study add?

- This study provides comprehensive census of human meniscus cells using single-cell RNA sequencing, and demonstrating CD146+ meniscus cells are stem/progenitor cells.
- Interleukin 1β induced activation of degenerated meniscus progenitor cells (DegP) is a potential mechanism contributing to meniscus degeneration.

How might this impact on clinical practice or future developments?

- CD146+ meniscus cells have potential in meniscus tissue engineering, and DegP could be a possible therapeutic target for meniscus degeneration.

INTRODUCTION

The menisci of mammals are crescent-shaped tissues, comprised a medial and a lateral component.[1] The meniscus plays an important role in joint stability, shock absorption, distribution of contact forces, joint lubrication and proprioception. The vascularisation of the meniscus decreased with ageing. The meniscus is fully vascularised during prenatal development and shortly after birth, however, only 10%–25% of mature meniscus contains blood vessels.[2] According to these differences in blood supply, the meniscus can be distinguished by the outer vascular region (red zone), inner avascular region (white zone) and the red-white zone between the red and white zones. The outer zone of the meniscus contains 90% type I collagen while the inner zone contains 60% type II collagen and 40% type I collagen.[3]

The cell types of the meniscus are heterogeneous, wherein the inner region contains chondrocyte-like morphology cells and the outer region fibroblast-like cells.[4] Recently, stem/progenitor cells were suggested to be present in the meniscus to promote meniscus injury repair.[5 6] Gamer *et al* isolated meniscus stem/progenitor cells by meniscus explant culture in vitro, characterising these cells with clonogenicity properties and abundantly expressed CD44 and Sca-1.[7] However, the cell-type composition and cell distribution in the menisci, as well as biochemical markers for meniscus stem cell/progenitor for use in tissue engineering, remain to be elucidated.

The relationship between meniscus degeneration and knee osteoarthritis (OA) is complex. Meniscus degenerative tears were found to be associated with increased cartilage loss in the same compartment, especially in posterior horn tears.[8 9] Fuller *et al* found that both inner and outer zone meniscal cells are responsive to the inflammatory cytokines IL-1α and TNF-α in an ovine *in vitro* model, which leads to cytokine-induced collagenolysis and aggrecanolysis.[10 11] These studies demonstrated the importance of meniscus degeneration in OA development and its contribution to joint disease in general. Degenerative meniscus accompanied by water content increased the wet weight, while collagen and total glycosaminoglycans (GAG) decreased.[12] However, the variety of cell types and corresponding biological markers, as well as the biological targets for the treatment for meniscus degeneration, have not yet been fully determined.

Single-cell RNA sequencing (scRNA-seq) is a well-established and powerful method to investigate transcriptomic cell-to-cell variation, which can be used to identify various cell types and provide insights into physiological and pathological processes.[13 14] Here, we used scRNA-seq to chart a comprehensive census of meniscus cells. We identified various cell subsets and their gene signatures to determine their differentiation relationships and characterise diversity within specific cell types. We also demonstrated the existence of meniscus stem/progenitor cells and their corresponding marker genes. Finally, we investigated the integral influence of meniscus degeneration on meniscus cellular heterogeneity and identified a potential therapeutic target.

MATERIALS AND METHODS

Isolation of human meniscus cells

Human meniscus tissues were dissected away from the synovium, and then cut into small pieces. Next, these small pieces were digested by 4 mg/mL protease (Roche 11459643001) for 1 hour and 2 mg/mL collagenase P (Roche 11213873001) for 6–10 hours.

RESULTS

scRNA-seq census of healthy human meniscus identified seven distinct cell populations

Our results showed that the cell quality was satisfactory for our single cell sequencing (online supplementary figure S1). To determine the cellular composition of human meniscus cells, we profiled meniscus cells from healthy human meniscus (n=3) using scRNA-seq. Unbiased clustering of the meniscus resulted in seven clusters originating from healthy human meniscus, including five empirically defined populations and two novel populations (figure 1A). Concretely, the following cells were identified: (1) endothelial cells (EC, expressing CD93 and CDH5),[15] (2) cartilage progenitor cells (CPC, expressing CDK1 and BIRC5),[16] (3) regulatory chondrocytes (RegC, expressing BMP2 and FOSL1),[17] (4) fibrochondrocytes (FC, expressing COL1A1, COL3A1 and COL6A1),[16 18] (5) prehypertrophic chondrocytes (PreHTC, expressing MMP1 and TNFAIP6),[19 20] (6) fibrochondrocyte progenitors (FCP, expressing both the fibrochondrocyte genes COL1A1 and COL3A1 and the mesenchymal stem cell marker genes MCAM and MYLK)[21] and (7) proliferate fibrochondrocytes (ProFC, expressing both the fibrochondrocyte gene COL1A1 and growth factors FGF7 and CTGF)[22] (figure 1B,C). FC and RegC were abundant, while FCP and EC were relatively rare.

To study the distribution of different cell clusters, we used immunohistochemistry to detect marker gene expression. MYLK, a marker gene of FCP, was mainly expressed on the meniscus surface, while the RegC gene marker BMP2 was mainly expressed in the middle of the meniscus. CD93 is the marker gene of EC and is mainly expressed around the vessels in the red zone, while the PreHTC marker ZIP8 was mainly expressed in white zone. No difference was found between the red and white areas regarding the expression of ProFC marker COL1A1, FC marker COL3A1 and CPC marker CDK1 (figure 1D).

Identification of population of human meniscus progenitor cells

To investigate the relationship between the different cell clusters, we used the Monocle method to reconstruct the pseudospace trajectory. We found that EC and FCP existed at the start of the pseudospace trajectory, and ProFC located in front of FC, while PreHTC was behind of FC. FC and CPC were distributed along the trajectory, and RegC was mainly distributed at the end (online supplementary figure S2A,B).

Since FCP existed at the start of the pseudospace trajectory, we investigated whether it had properties characteristic of a progenitor. Pathway analysis showed that pathways involved in focal adhesion, extracellular matrix (ECM)–receptor interaction and TGFβ signalling were activated (figure 2A). FCP expressed the mesenchymal stem cell marker MCAM (CD146) (figure 2B), as well as classical markers of myofibroblasts, including ACTA2, MYLK and MYL9 (online supplementary figure S2C).

We isolated the CD146+ primary human meniscus cells using fluorescence-activated cell sorting (FACS) and found that the proportion of CD146+ meniscus cells was near 2.7% (figure 2C). CD146+ cells had the ability to differentiate into various cell lineages, including osteoblasts and adipocytes (figure 2D). Next, we examined the clonogenicity of CD146+ cells. A total of 2000 CD146+ and CD146– cells were seeded in 12-well plates and cultured for 7 days. After culturing, the number of colonies in the CD146+ group was significantly higher than that of the colonies in the CD146– cells group (figure 2E). The fact that the cells in the CD146– group were able to form colonies suggests that another cell cluster may also have progenitor properties. CD93 is the specific marker of ECs, so we used FACS to obtain CD146+/CD93+ meniscus cells (EC) and CD146+/CD93– meniscus cells (FCP). Further experiments showed that these two clusters have progenitor properties (online supplementary figure S2D).

Single-cell trajectory branch points correspond to FCP differentiation

To study the differentiation of FCP into subset clusters and the corresponding gene expression, we selected FCP, ProFC, FC, PreHTC and RegC to construct a new trajectory containing two termini corresponding to two distinct cell fates (figure 3A). The root of the trajectory was mainly populated by FCP and ProFC, while the two termini of the tree were populated by FC and PreHTC for fate 1, RegC and PreHTC for fate 2 (figure 3B). Next, we assessed the expression of genes regulated during FCP differentiation in cells at fates 1 and 2 of the trajectory. The expression of MYLK, CNN1, FGF7 and COL1A1 were found to be similar. While MYLK and CNN1 expression was markedly reduced from the root through to both fates, FGF7 and COL1A1 expression was upregulated early in FCP differentiation and downregulated in cells differentiating into both fates (figure 3B,C). ADAMTS4 and MMP1 were slightly upregulated at early stage differentiation, and notably upregulated in cells at fate 1 and downregulated at fate 2. On the contrary, FOSL1 and

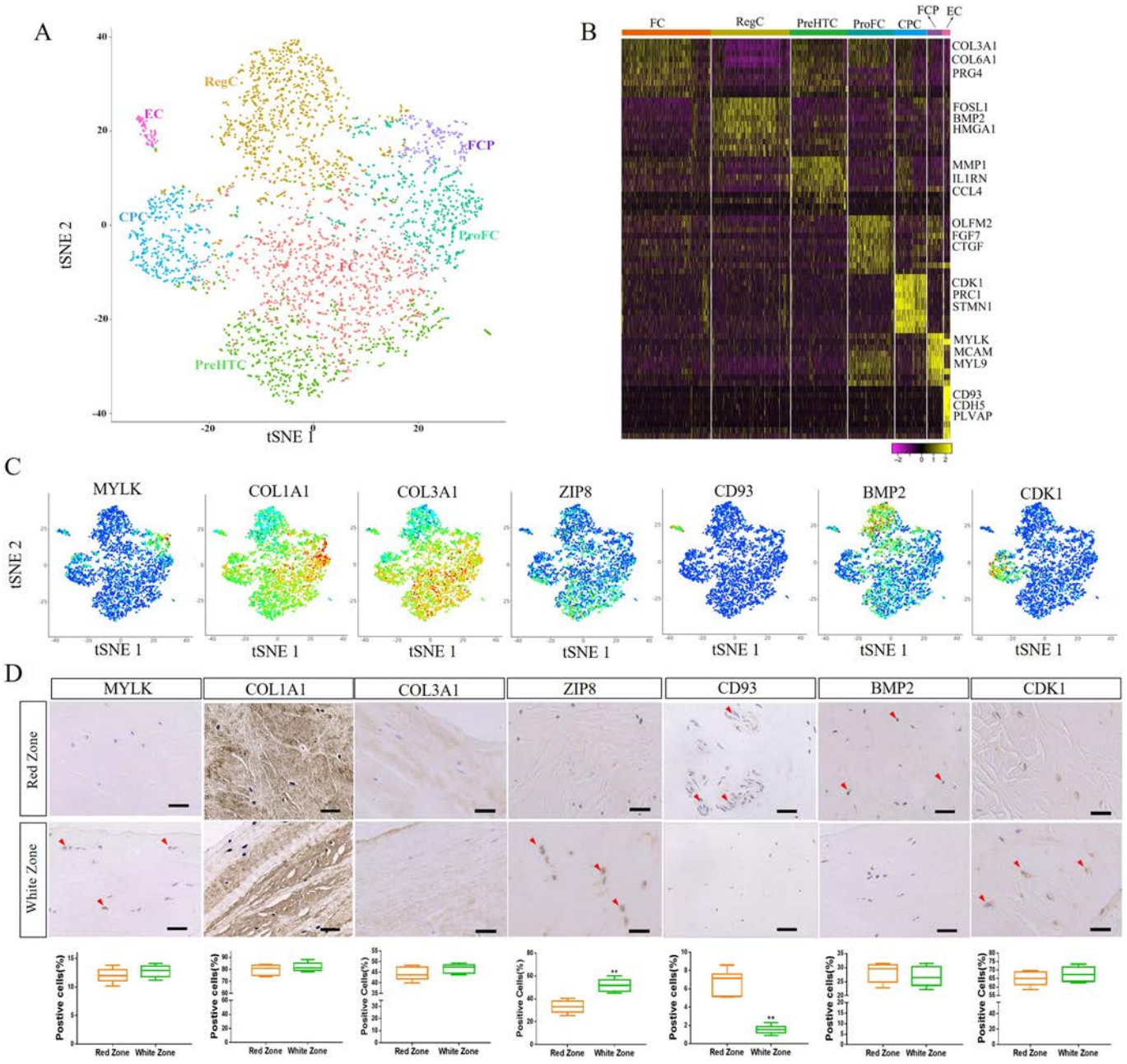


Figure 1 A single-cell atlas of healthy human meniscus. (A) Seven healthy human meniscus cell clusters. t-Distributed stochastic neighbour embedding (t-SNE) of 3639 cells (mixed with cell fractions, n=3), annotated post-hoc and coloured by clustering. (B) Heatmap revealing the scaled expression of differentially expressed genes for each cluster. (C) Dot plots showing the expression of the indicated markers for each cell cluster on the t-SNE map. (D) Representative immunohistochemistry staining of MYLK, COL1A1, COL3A1, ZIP8, CD93, BMP2 and CDK1 in white and red zones of healthy human meniscus tissues, and quantification of positive cells displayed by box plot (n=6). Scale bar, 50 µm. **p<0.01. CPC, cartilage progenitor cells; EC, endothelial cells; FC, fibrochondrocytes; FCP, fibrochondrocyte progenitors; PreHTC, prehypertrophic chondrocytes; ProFC, proliferate fibrochondrocytes; RegC, regulatory chondrocytes.

BMP2 expression slightly decreased at cells from the root to fate 1, but markedly increased in cells differentiating via fate 2.

To confirm the single-cell trajectory, we analysed the meniscus developmental process in vivo and studied the expression of marker genes in the mice meniscus at 1, 2, 3, 4, 8, 26 and 52 weeks. COL1A1 expression increased gradually after birth and peaked at 4 weeks, then decreased gradually with increasing age (figure 3E). MYLK expression decreased significantly with increasing age after 3 weeks (figure 3E). These expression patterns were consistent with two different fates of the trajectory, indicating that our scRNA-seq analysis correlated with the meniscus developmental process.

Systemic comparison of the single cell landscape between healthy human meniscus and degenerated meniscus

To comprehensively assess the changes in the human meniscus during degeneration, we first evaluated the histological changes in degenerated meniscus. The healthy meniscus was negative for Safranine O staining, while the degenerative meniscus was positive for staining (online supplementary figure S3). In addition,

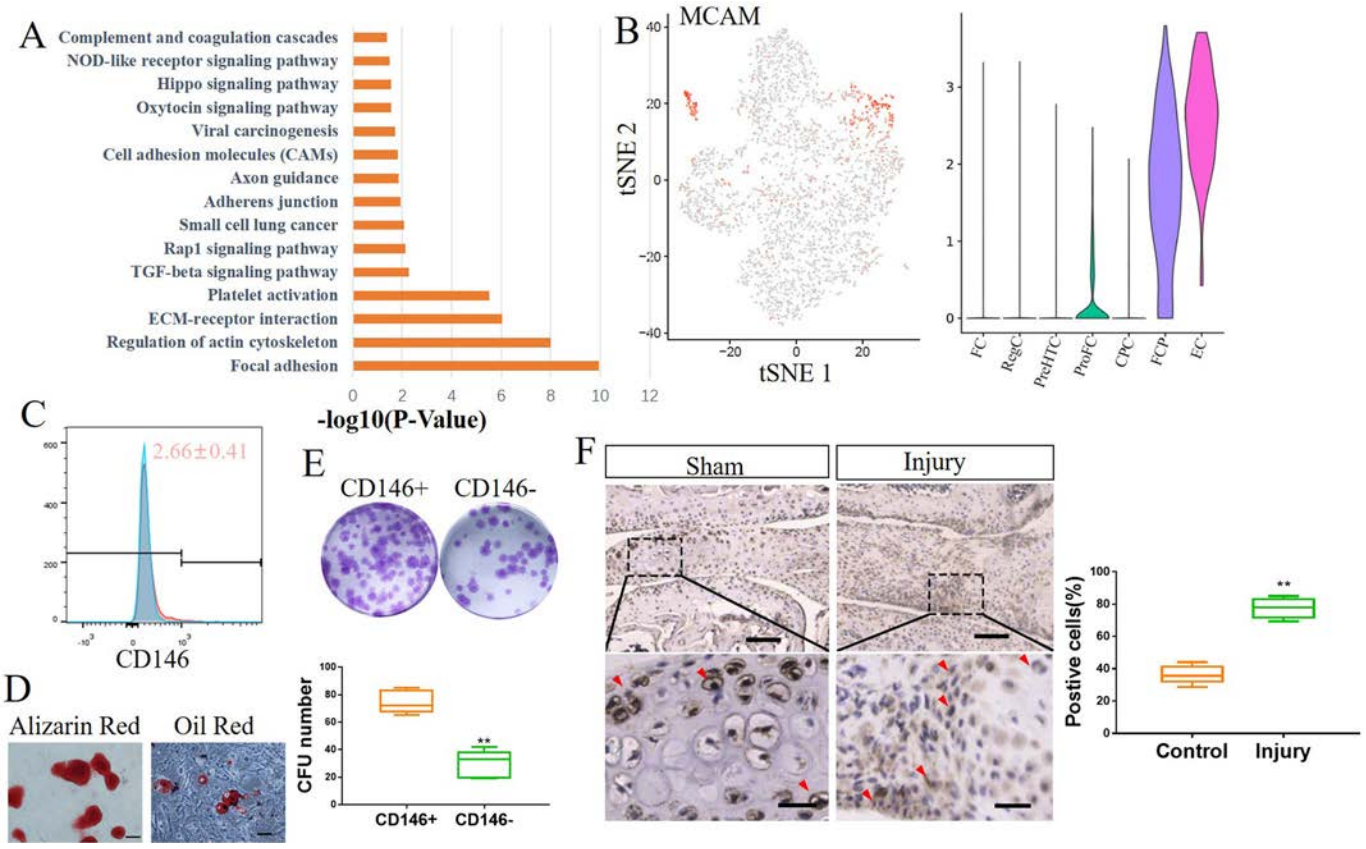


Figure 2 Identification of human meniscus progenitor cells. (A) The 15 most upregulated signal pathways in FCP. (B) Dot plots showing the MCAM expression on t-distributed stochastic neighbour embedding (t-SNE) map and Vin plot. (C) CD146 expression in healthy human meniscus cells determined by flow cytometry (mean±SD; n=3). (D) Alizarin red staining and oil red staining for CD146+ meniscus cells induced to osteogenic differentiation or adipogenic differentiation, respectively (n=5). Scale bar, 50 µm. (E) Colony-forming analysis of CD146+ and CD146– healthy human meniscus cells and quantification. n=5, **p<0.01. (F) IHC staining of MYLK in mice meniscus injury model, and quantification of positive cells. Scale bars, 200 µm (top) and 50 µm (bottom). n≥6, **p<0.01. CFU colony forming unit; CPC, cartilage progenitor cells; EC, endothelial cells; FC, fibrochondrocytes; FCP, fibrochondrocyte progenitors; NOD, nucleotide-binding oligomerisation domain; PreHTC, prehypertrophic chondrocytes; ProFC, proliferate fibrochondrocytes; RegC, regulatory chondrocytes.

the collagen fibre structure on the degenerated meniscus was disorganised (Figure 4A and online supplementary figure S4). Next, we compared the scRNA-seq between healthy meniscus and degenerated meniscus (figure 4B–E). As a result, we detected significant changes in the proportions of degenerated meniscus cell clusters, including three new clusters: (1) monocyte-derived dendritic cells (MoDC, expressing CD14 and S100A9),[23 24] (2) hypertrophic chondrocytes (HTC, expressing CCL20 and EREG)[25 26] and (3) degenerated meniscus progenitor cells (DegP), which are found in degenerated meniscus and express skeletal stem cell marker, such as GREM1[27] (figure 4C–G). Moreover, the proportion of EC and FCP expression was found to decrease in degenerated meniscus (figure 4E).

Alignment of single-cell trajectories indicates DegP is a key element for meniscus degeneration

CDCP1 (CD318) is highly expressed in DegP (online supplementary figure S4A). As such, we isolated the CD318+ primary human degenerated meniscus cells by FACS to verify the progenitor capacity. CD318+ cells were found to form colonies and differentiate into various cell lineages (online supplementary figure S4B,C), wherein DegP was a special population with progenitor characteristics, and was mainly found in the degenerative meniscus.

Next, we selected four clusters with progenitor properties, including FCP, ProFC, CPC and DegP, to construct a new trajectory. The trajectory's root was mainly populated by FCP and ProFC, while the two primary termini of the tree were populated by DegP and CPC for fate 1, and CPC for fate 2 (figure 5A,B). Although MCAM and MYLK were highly expressed at the root of the trajectory, their expression was markedly reduced along the root through to both fates 1 and 2 (figure 5C,D). BIRC5 and CDK1 were highly expressed at the end of fate 2, while GAS1, RAB3B and CDCP1 were highly expressed at the end of fate 1 (figure 5C,D and online supplementary figure S5A). However, in normal FCP differentiation, the expression of GAS1, RAB3B and CDCP1 was markedly reduced while progressing along from the root to both fates 1 and 2 (compared with figure 3, online supplementary figure S5B), indicating that fate 1 may be an aberrant cellular state during the degeneration process in meniscus.

We also verified the expression of marker genes by IHC staining. MYLK, an FCP marker gene, was downregulated in degenerated meniscus, while the DegP marker genes GAS1 and DNER were upregulated in degenerated meniscus, especially in areas where meniscus lesions were accompanied by cell proliferation (figure 5E).

Proinflammatory mediators, such as IL-1β, appeared to directly influence the degradative processes in the meniscus.[28 29]

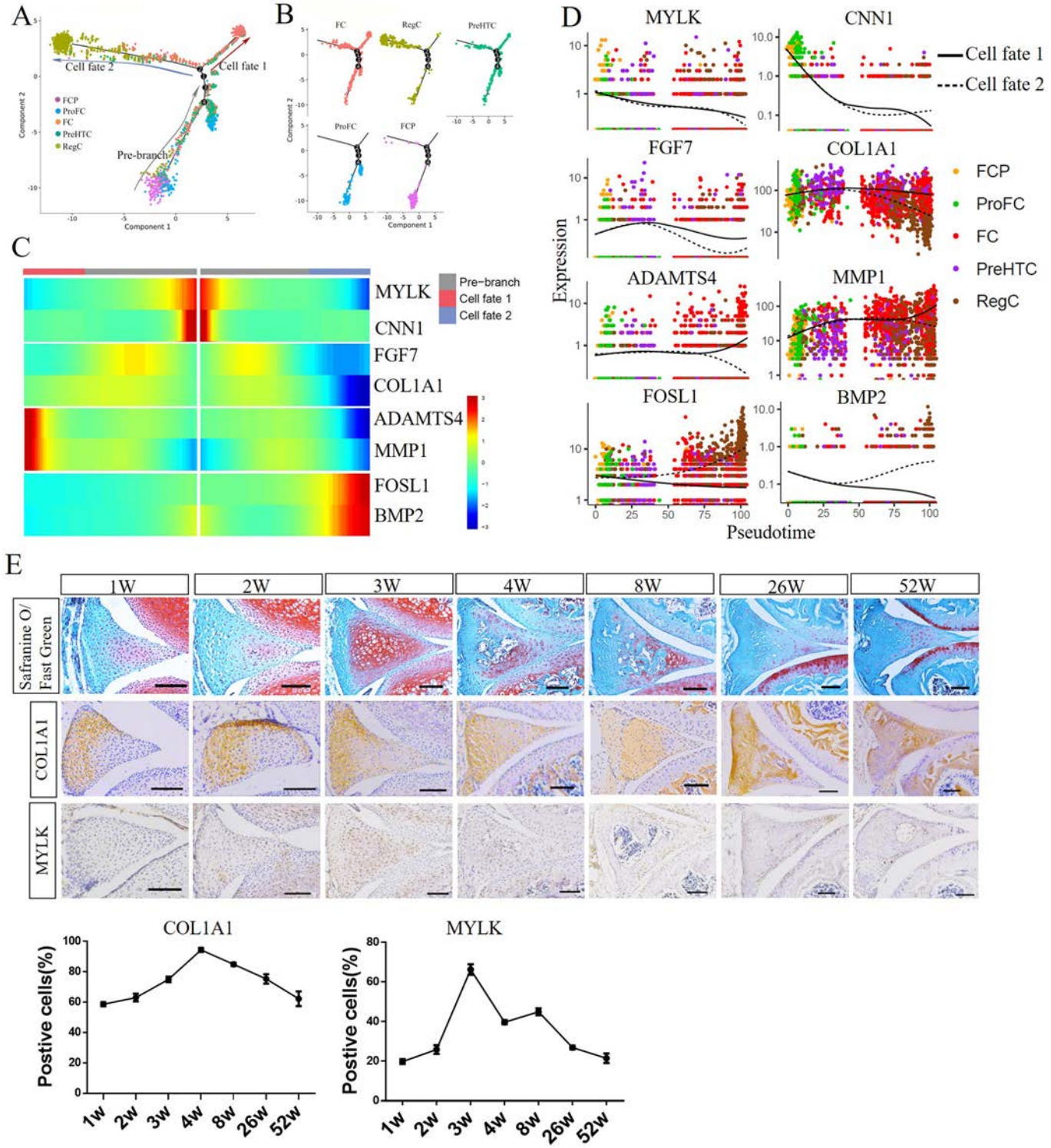


Figure 3 Single-cell trajectory branch points demonstrating FCP differentiation. (A, B) Monocle pseudotime trajectory showing the progression of FCP, ProFC, FC, PreHTC and RegC. (C) The expression of the genes in a branch-dependent manner. Each row indicates the standardised kinetic curves of a gene. The centre of the heatmap shows the kinetic curve value at the root of the trajectory. From the centre to the left of the heatmap, the kinetic curve progresses from the root along the trajectory to fate 1. Starting from the right, the curve from the root to fate 2. (D) Pseudotime kinetics of indicated genes from the root of the trajectory to fate 1 (solid line) and the cells up to fate 2 (dashed line). (E) Safranine O/Fast Green staining and immunohistochemistry staining of COL1A1 and MYLK in mice anterior meniscus at 1, 2, 3, 4, 8, 26 and 52 weeks, and quantification of positive cells (n≥3). Scale bar, 100 µm. CPC, cartilage progenitor cells; EC, endothelial cells; FC, fibrochondrocytes; FCP, fibrochondrocyte progenitors; PreHTC, prehypertrophic chondrocytes; ProFC, proliferate fibrochondrocytes; RegC, regulatory chondrocytes.

Therefore, we used IL-1β (5 ng/mL) to stimulate healthy human meniscus cells for 48 and 96 hours to detect any changes in CD146+ cells and CD318+ cells. IL-1β stimulation led to a significant reduction in CD146+ cells with an increasing stimulation time, while CD318+ cells significantly increased (figure 5F). We also used IL-1β to stimulate degenerated

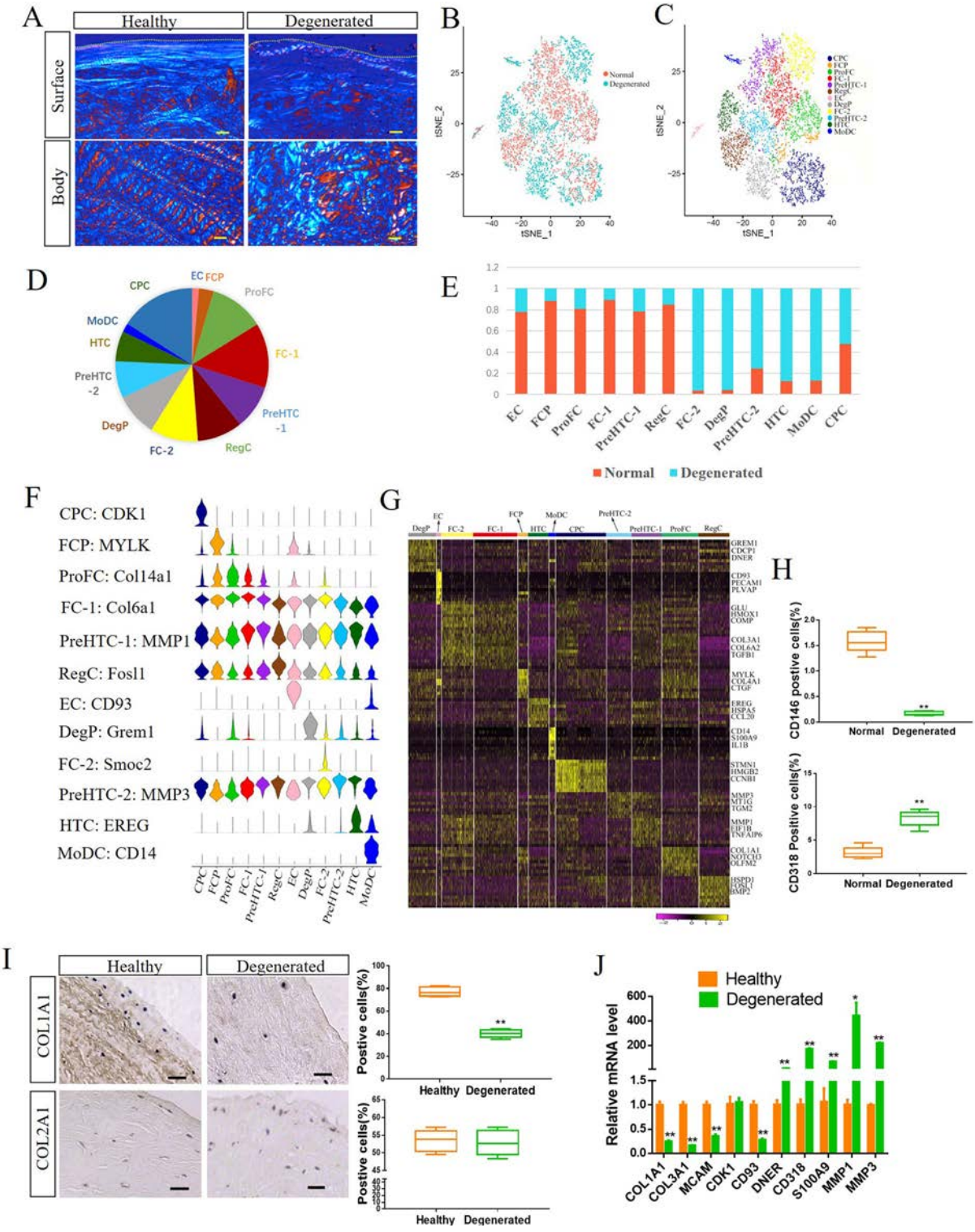


Figure 4 Comparison of the single cell landscape between healthy human meniscus and degenerated meniscus. (A) Representative polarised light microscopy images of healthy human and degenerated meniscus. The white and red colours in the angle images are 90° apart in orientation. Dashed lines indicate the surface of the meniscus. Scale bar, 100 µm. (B) Merged t-distributed stochastic neighbour embedding (t-SNE) of single-cell RNA sequencing of healthy meniscus cells and degenerated meniscus cells. (C) Twelve healthy human and degenerated meniscus cell clusters at t-SNE. (D) Proportion of each cluster to the total cells. (E) Proportion of healthy and degenerated meniscus cells in each cluster. (F) Expression of representative marker genes in Vin plot. (G) Heatmap revealing the scaled expression of differentially expressed genes for each cluster. (H) CD146 and CD318 expression in healthy human meniscus cells and degenerated meniscus cells determined by flow cytometry. n≥5, **p<0.01. (I) Representative IHC staining of COL1A1 and COL2A1 healthy human meniscus and degenerated meniscus, and quantification of positive cells. Scale bar, 50 µm. n≥5, **p<0.01. (J) The expression of indicated marker genes in human healthy meniscus cells and degenerated meniscus cells were detected by qRT-PCR. *p<0.05, **p<0.01, otherwise, not significant. n=3, *p<0.05, **p<0.01. CPC, cartilage progenitor cells; DegP, degenerated meniscus progenitor cells; EC, endothelial cells; FC, fibrochondrocytes; FCP, fibrochondrocyte progenitors; HTC, hypertrophic chondrocytes; MoDC, monocyte-derived dendritic cells; PreHTC, prehypertrophic chondrocytes; ProFC, proliferate fibrochondrocytes; RegC, regulatory chondrocytes.

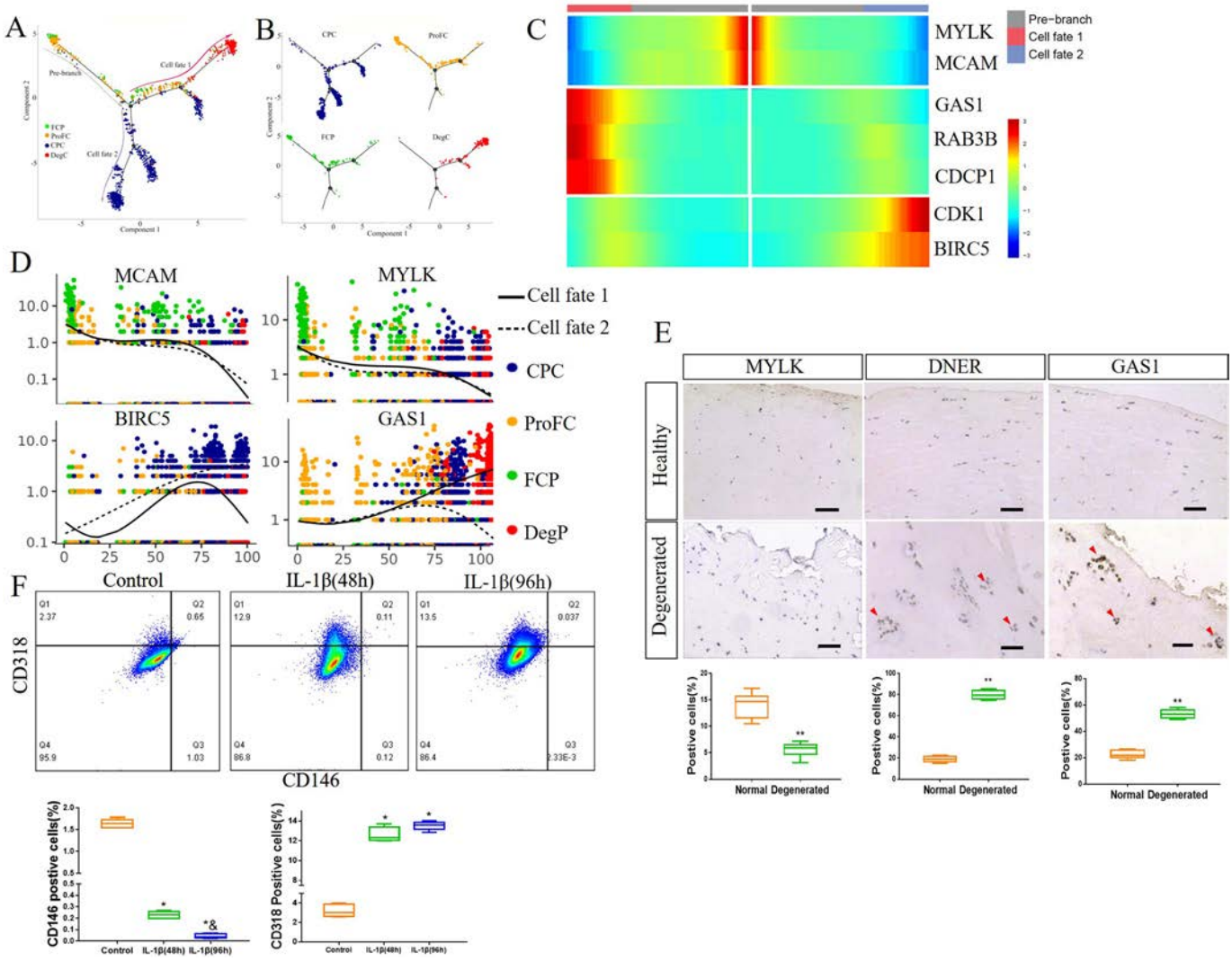


Figure 5 Identification of degenerated meniscus progenitor cells (DegP) as a key element for meniscus degeneration. (A, B) Monocle pseudotime trajectory showing the progression of FCP, ProFC, CPC and DegP. (C) From the centre to the left of the heatmap, the kinetic curve from the root along the trajectory to fate 1. Starting from the right, the curve from the root to fate 2. FCP markers MYLK and MCAM, DegP markers GAS1, Rab3B and CDCP1 and CPC markers CDK1 and BIRC5 expressed from the root to each branch. (D) Pseudotime kinetics of indicated genes from the root of the trajectory to fate 1 (solid line) and the cells up to fate 2 (dashed line). (E) Representative IHC staining of MYLK, GAS1 and DNER in healthy human meniscus and degenerated meniscus, and quantification of positive cells. Scale bar, 50 µm. n=6, **p<0.01. (F) Healthy human meniscus cells were treated with 5 ng/mL IL-1β for 48 hours or 96 hours. Phosphate buffer saline (PBS) was used as a negative control. CD146 and CD318 expression was determined by flow cytometry. n≥5, * versus control, p<0.05; & versus IL-1β (48 hours), p<0.05. CPC, cartilage progenitor cells; DeP, degenerated meniscus progenitor cell; FCP, fibrochondrocyte progenitors; ProFC, proliferate fibrochondrocytes.

human meniscus cells and get similar results (online supplementary figure S5C). These results suggested that the decrease of CD146+ cells and the increase of CD318+ cells caused by various pathogenic factors such as IL-1β, may be an important mechanism of meniscus degeneration.

Activation of TGFβ signalling pathway attenuates the increase in CD318+ cells in degenerated meniscus

Previous studies have shown that the activation of TGFβ signalling enhances the differentiation ability of meniscus progenitors.[30 31] Our scRNA-seq analysis and IHC staining showed that TGFβ1, a ligand of the transforming growth factor-β (TGFβ) signalling pathway, was highly expressed in healthy meniscus cells (figure 6A,B). We also compared the differences in gene expression between FC-1 and FC-2, PreHTC-1 and PreHTC-2, two cell types found in both healthy and degenerated meniscus. Compared with the clusters mainly found in degenerated meniscus (FC-2 and PreHTC-2), the clusters found in healthy meniscus (FC-1 and PreHTC-1) were upregulated by the TGFβ signalling pathway (online supplementary figure S6) and highly expressed COL1A1, COL3A1 and TGFβ1 (figure 6C,D). Next, we investigated the effect of TGFβ1 on DegP. Primary human degenerated meniscus cells were treated with 5 ng/mL TGFβ1 for 48 hours or 96 hours. Flow cytometry demonstrated that TGFβ1 treatment significantly reduced the number of CD318+ cells in a time depend manner (figure 6E), and qRT-PCR showed TGFβ1 treatment significantly increasing COL1A1, COL3A1 and CDK1 expression while decreasing CD318, S100A9, MMP1 and MMP3 expression (figure 6F), indicating that TGFβ1 may be able to delay the degeneration of meniscus.

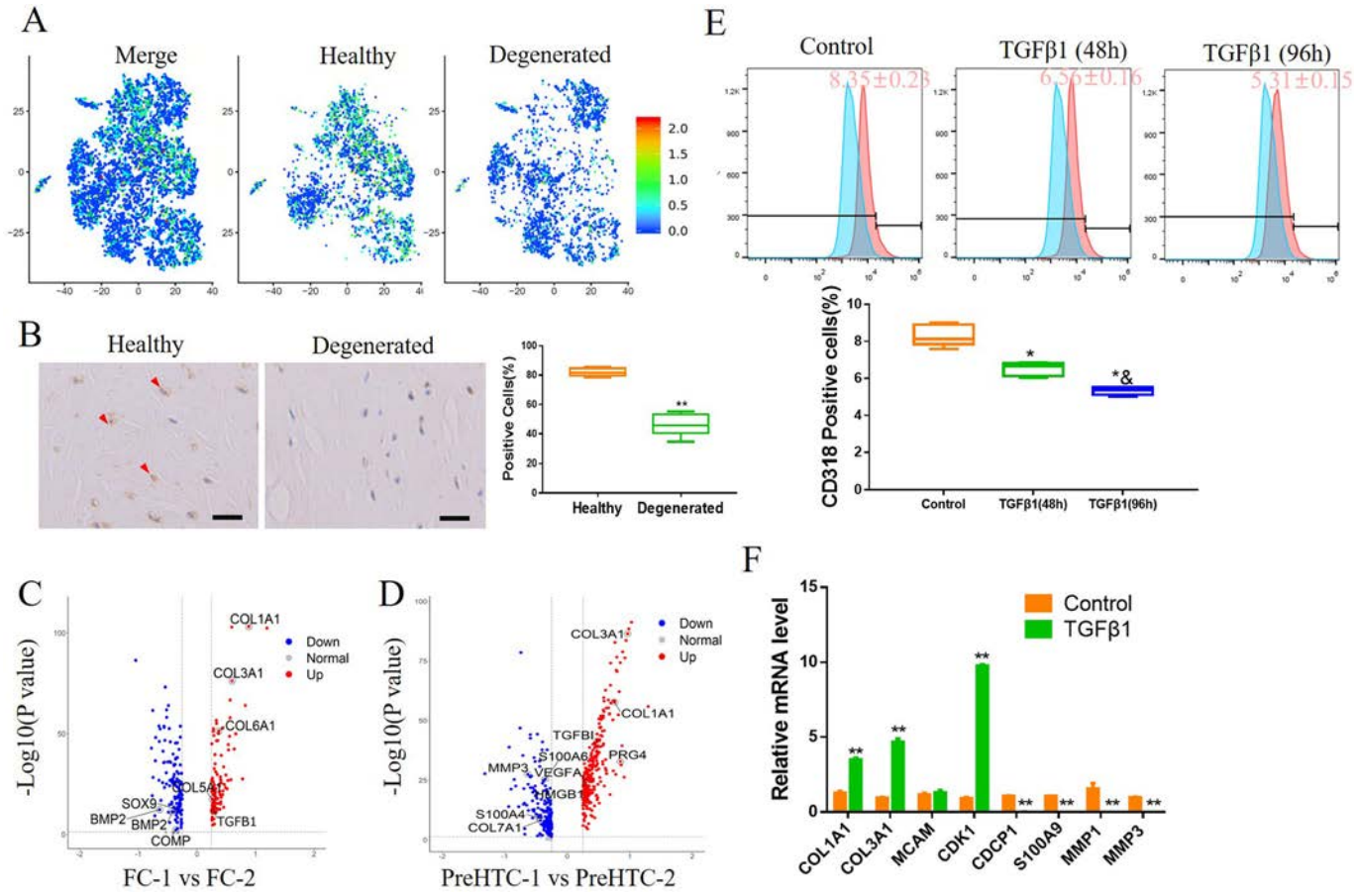


Figure 6 Activation of TGFβ signalling pathway attenuates the increase in CD318+ cells in degenerated meniscus. (A) The expression of TGFβ1 on merged and split t-distributed stochastic neighbourembedding map. (B) IHC staining of TGFβ1 on human healthy meniscus and degenerated meniscus. n=6, **p<0.01. (C) Volcano plot comparing the gene expression between FC-1 and FC-2. Each plot represents one gene. (D) Volcano plot comparing the gene expression between PreHTC-1 and PreHTC-2. Each plot represents one gene. (E) Human degenerated meniscus cells were treated with 5 ng/mL TGFβ1 for 48 hours or 96 hours. PBS was used as a negative control. CD318 expression was determined by flow cytometry (n≥5). * vs control, p<0.05; & vs TGFβ1 (48 hours), p<0.05. (F) Human degenerated meniscus cells were treated with 5 ng/mL TGFβ1 or PBS as negative control. The expression of indicated marker genes were detected by qRT-PCR. n=3, **p<0.01.

DISCUSSION

An increasing number of studies are supporting the idea that cell-based strategies effectively improve meniscus repair and regeneration.[32 33] However, it is still not clear which cell type is most effective for meniscus repair.[34 35] Recently, meniscus stem/progenitor cells have been considered as the most suitable cell type for meniscus injury repair due to them having the same tissue origin and histocompatibility,[36 37] however, the characteristics, marker genes and isolation methods of human meniscus progenitor cells have not yet been fully elucidated. Gamer *et al* isolated meniscus progenitor cells from mice meniscus grown in explant cultures, and carried out flow cytometry analysis to show that these cells highly expressed CD44 and Sca-1.[7] Shen *et al* digested human meniscus using collagenase and seeded the cells at a low density to form colonies. The subsequent flow cytometry analysis showed that these cell highly expressed CD90 (THY1) and CD105 (ENG), and the intra-articular injection of these cells promoted rat meniscus regeneration and ameliorated OA.[37] Our scRNA-seq results also show the high expression of CD90 and CD105 in FCP, however, they were also highly expressed in FC-1 and FC-2. Thus, CD90 and CD105 were not markers specific to meniscus progenitor cells.

In our scRNA-seq results, EC was found to exist at the start of the pseudospace trajectory, which plays an important role in the development, degeneration and repair of the meniscus.[38 39] Miller and Rydell isolated meniscus EC for the first time in 1993, and proved these cells had the ability to self-renew and maintain their characteristics after 10 passages.[40] EC is not only able to generate vessels to maintain blood supply, but also promote the migration of meniscus cells. Yuan *et al* found that EC could enhance meniscus cell migration by activating endothelin signalling.[41] Notably, we identified CD146 specifically expressed in EC and FCP, suggesting that CD146+ cells can be used in cell-based scaffolding for meniscus injury repair, and may also be a target for recruitment of meniscus progenitor cells by growth factors to participate in meniscus injury repair in cell-free strategies.

We identified three cell clusters specific to degenerated meniscus, including MoDC, HTC and DegP, where DegP is a novel cluster and has the characteristics of progenitor cells. Our scRNA-seq demonstrated that the expression of DegP markers, such as GAS1, RAB3B and CD318, increased rapidly at the end of the differentiation of FCP to DegP, which was contrary to the normal differentiation procedure, suggesting that this differentiation process was the result of an aberrant cellular state. IHC staining also showed that GAS1 and RAB3B were highly expressed in meniscus with severe lesions, which was accompanied by cell proliferation. CD318 has been previously demonstrated to be highly expressed in haematopoietic progenitors[42]

and muscle progenitors.[43] Iwata *et al* revealed that CD318 is a CD146 negative subset of bone marrow fibroblasts and regulates cytokine expression.[44] Previous studies have shown that inflammatory cytokines such as IL-1β and TNF-α induce meniscus metabolic responses and result in degeneration.[28 45] In our study, IL-1β was used to induce the inflammatory response in human meniscus cells. We demonstrated that IL-1β decreased CD146+ cells and increased CD318+ cells in both healthy and degenerated meniscus cells. These results demonstrate that DegP plays a crucial role in meniscus degeneration and may be used as a marker to evaluate meniscus degeneration or as a target for the treatment of meniscus degeneration.

TGFβ is widely used in meniscus tissue engineering, owing to its promotion of meniscus injury repair and regeneration through the promotion of fibrochondrocyte proliferation and recruitment of meniscus progenitor cells.[30 46 47] TGFβ also regulates the meniscus degeneration process, while the postnatal deletion of TGFβ signalling reporter ALK5 accelerates meniscus degeneration.[31] Our scRNA-seq results showed that TGFβ1 is highly expressed in FC-1 and FC-2, and that its overall expression in degenerated meniscus is decreased. Treatment with TGFβ1 has been previously found to enhance the mechanical properties of tissue-engineered fibrocartilage.[48] Our results revealed that TGFβ1 attenuated the proportion of CD318+ cells in human degenerated meniscus, suggesting that TGFβ1 may be used to suppress meniscus degeneration.

In conclusion, our scRNA-seq results provided a clearer and more consistent definition of the cellular components of human meniscus, and the ways in which specific clusters contribute to meniscus development and aberrant degeneration. Our analysis identified the meniscus progenitors with potential in meniscus tissue engineering. We also demonstrated an important mechanism of meniscus degeneration and provided experimental evidence for a therapeutic strategy.

Acknowledgements We thank Yujie Liu and Bo Zhang from NovelBioinformatics for the support of bioinformatics analysis.

Contributors All listed authors meet the criteria for authorship and have contributed to the study design, data generation, data analysis, manuscript writing and manuscript review.

Funding This study was funded by the National Natural Science Foundation of China (no. 81874016, 81672145, 81472101, 81572119) and the Science and Technology Project of Guangzhou City, China (201710010164).

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval Protocol approved by the Ethical Committee of The First Affiliated Hospital of Sun Yat-sen University.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. The single-cell RNA-seq data, quality control information and cluster information are available at the NCBI's Gene Expression Omnibus (GEO) data repository with the accession ID GSE133449.

ORCID iDs
Hao Sun https://orcid.org/0000-0001-8829-0468
Zhiqi Zhang https://orcid.org/0000-0001-9086-2453

REFERENCES

1 Makris EA, Hadidi P, Athanasiou KA. The knee meniscus: Structure–function, pathophysiology, current repair techniques, and prospects for regeneration. *Biomaterials* 2011;32:7411–31.
2 Clark CR, Ogden JA. Development of the menisci of the human knee joint. Morphological changes and their potential role in childhood meniscal injury. *J Bone Joint Surg Am* 1983;65:538–47.
3 Danso EK, Oinas JMT, Saarakkala S, *et al*. Structure-Function relationships of human meniscus. *J Mech Behav Biomed Mater* 2017;67:51–60.
4 Cheung HS. Distribution of type I, II, III and V in the pepsin solubilized collagens in bovine menisci. *Connect Tissue Res* 1987;16:343–56.
5 Huang H, Wang S, Gui J, *et al*. A study to identify and characterize the stem/progenitor cell in rabbit meniscus. *Cytotechnology* 2016;68:2083–103.
6 Seol D, Zhou C, Brouillette MJ, *et al*. Characteristics of meniscus progenitor cells migrated from injured meniscus. *J Orthop Res* 2017;35:1966–72.
7 Gamer LW, Shi RR, Gendelman A, *et al*. Identification and characterization of adult mouse meniscus stem/progenitor cells. *Connect Tissue Res* 2017;58:238–45.
8 Crema MD, Guermazi A, Li L, *et al*. The association of prevalent medial meniscal pathology with cartilage loss in the medial tibiofemoral compartment over a 2-year period. *Osteoarthritis Cartilage* 2010;18:336–43.
9 Sihvonen R, Paavola M, Malmivaara A, *et al*. Arthroscopic partial meniscectomy versus placebo surgery for a degenerative meniscus tear: a 2-year follow-up of the randomised controlled trial. *Ann Rheum Dis* 2018;77:188–95.
10 Fuller ES, Smith MM, Little CB, *et al*. Zonal differences in meniscus matrix turnover and cytokine response. *Osteoarthritis Cartilage* 2012;20:49–59.
11 Fuller E, Little CB, Melrose J. Interleukin-1α induces focal degradation of biglycan and tissue degeneration in an in-vitro ovine meniscal model. *Exp Mol Pathol* 2016;101:214–20.
12 Herwig J, Egner E, Buddecke E. Chemical changes of human knee joint menisci in various stages of degeneration. *Ann Rheum Dis* 1984;43:635–40.
13 Svensson V, Natarajan KN, Ly L-H, *et al*. Power analysis of single-cell RNA-sequencing experiments. *Nat Methods* 2017;14:381–7.
14 Tian L, Dong X, Freytag S, *et al*. Benchmarking single cell RNA-sequencing analysis pipelines using mixture control experiments. *Nat Methods* 2019;16:479–87.
15 Baryawno N, Przybylski D, Kowalczyk MS, *et al*. A cellular taxonomy of the bone marrow stroma in homeostasis and leukemia. *Cell* 2019;177:1915–32.
16 Ji Q, Zheng Y, Zhang G, *et al*. Single-cell RNA-seq analysis reveals the progression of human osteoarthritis. *Ann Rheum Dis* 2019;78:100–10.
17 Dunn SL, Soul J, Anand S, *et al*. Gene expression changes in damaged osteoarthritic cartilage identify a signature of non-chondrogenic and mechanical responses. *Osteoarthritis Cartilage* 2016;24:1431–40.
18 Gunja NJ, Athanasiou KA. Passage and reversal effects on gene expression of bovine meniscal fibrochondrocytes. *Arthritis Res Ther* 2007;9.
19 Baker J, Falconer AMD, Wilkinson DJ, *et al*. Protein kinase D3 modulates MMP1 and MMP13 expression in human chondrocytes. *PLoS One* 2018;13:e0195864.
20 Won Y, Shin Y, Chun C-H, *et al*. Pleiotropic roles of metallothioneins as regulators of chondrocyte apoptosis and catabolic and anabolic pathways during osteoarthritis pathogenesis. *Ann Rheum Dis* 2016;75:2045–52.
21 Puram SV, Tirosh I, Parikh AS, *et al*. Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell* 2017;171:1611–24.
22 Berent-Maoz B, Montecino-Rodriguez E, Signer RAJ, *et al*. Fibroblast growth factor-7 partially reverses murine thymocyte progenitor aging by repression of INK4A. *Blood* 2012;119:5715–21.
23 Loh C, Park S-H, Lee A, *et al*. Tnf-Induced inflammatory genes escape repression in fibroblast-like synoviocytes: transcriptomic and epigenomic analysis. *Ann Rheum Dis* 2019;78:1205–14.
24 McGovern N, Schlitzer A, Gunawan M, *et al*. Human dermal CD14+ cells are a transient population of monocyte-derived macrophages. *Immunity* 2014;41:465–77.
25 Alaaeddine N, Antoniou J, Moussa M, *et al*. The chemokine CCL20 induces proinflammatory and matrix degradative responses in cartilage. *Inflamm Res* 2015;64:721–31.
26 Chen Y-J, Chang W-A, Wu L-Y, *et al*. Systematic analysis of transcriptomic profile of chondrocytes in osteoarthritic knee using next-generation sequencing and bioinformatics. *J Clin Med* 2018;7:E535.
27 Worthley DL, Churchill M, Compton JT, *et al*. Gremlin 1 identifies a skeletal stem cell with bone, cartilage, and reticular stromal potential. *Cell* 2015;160:269–84.
28 Cook AE, Cook JL, Stoker AM. Metabolic responses of meniscus to IL-1β. *J Knee Surg* 2018;31:834–40.
29 McNulty AL, Rothfusz NE, Leddy HA, *et al*. Synovial fluid concentrations and relative potency of interleukin-1 alpha and beta in cartilage and meniscus degradation. *J Orthop Res* 2013;31:1039–45.
30 Muhammad H, Schminke B, Bode C, *et al*. Human migratory meniscus progenitor cells are controlled via the TGF-β pathway. *Stem Cell Reports* 2014;3:789–803.
31 Wang Q, Tan Q, Xu W, *et al*. Postnatal deletion of *Alk5* gene in meniscal cartilage accelerates age-dependent meniscal degeneration in mice. *J Cell Physiol* 2019;234:595–605.
32 Shimomura K, Rothrauff BB, Hart DA, *et al*. Enhanced repair of meniscal hoop structure injuries using an aligned electrospun nanofibrous scaffold combined with a mesenchymal stem cell-derived tissue engineered construct. *Biomaterials* 2019;192:346–54.

33 Zhang Z-Z, Zhou Y-F, Li W-P, *et al*. Local administration of magnesium promotes meniscal healing through homing of endogenous stem cells: a proof-of-concept study. *Am J Sports Med* 2019;47:954–67.
34 Moradi L, Vasei M, Dehghan MM, *et al*. Regeneration of meniscus tissue using adipose mesenchymal stem cells-chondrocytes co-culture on a hybrid scaffold: In vivo study. *Biomaterials* 2017;126:18–30.
35 Zellner J, Pattappa G, Koch M, *et al*. Autologous mesenchymal stem cells or meniscal cells: what is the best cell source for regenerative meniscus treatment in an early osteoarthritis situation? *Stem Cell Res Ther* 2017;8.
36 Gui J, Zhang J, Huang H. Isolation and characterization of meniscus derived stem cells from rabbit as a possible treatment for damaged meniscus. *Curr Stem Cell Res Ther* 2015;10:353–63.
37 Shen W, Chen J, Zhu T, *et al*. Intra-Articular injection of human meniscus stem/progenitor cells promotes meniscus regeneration and ameliorates osteoarthritis through stromal cell-derived factor-1/CXCR4-mediated homing. *Stem Cells Transl Med* 2014;3:387–94.
38 Ashraf S, Wibberley H, Mapp PI, *et al*. Increased vascular penetration and nerve growth in the meniscus: a potential source of pain in osteoarthritis. *Ann Rheum Dis* 2011;70:523–9.
39 Rai MF, Patra D, Sandell LJ, *et al*. Transcriptome analysis of injured human meniscus reveals a distinct phenotype of meniscus degeneration with aging. *Arthritis Rheum* 2013;65:2090–101.
40 Miller RR, Rydell PA. Primary culture of microvascular endothelial cells from canine meniscus. *J Orthop Res* 1993;11:907–11.
41 Yuan X, Eng GM, Arkonac DE, *et al*. Endothelial cells enhance the migration of bovine meniscus cells. *Arthritis Rheumatol* 2015;67:182–92.
42 Takeda H, Fujimori Y, Kai S, *et al*. CD318/CUB-domain-containing protein 1 expression on cord blood hematopoietic progenitors. *Exp Ther Med* 2010;1:497–501.
43 Uezumi A, Nakatani M, Ikemoto-Uezumi M, *et al*. Cell-surface protein profiling identifies distinctive markers of progenitor cells in human skeletal muscle. *Stem Cell Reports* 2016;7:263–78.
44 Iwata M, Torok-Storb B, Wayner EA, *et al*. Cdcp1 identifies a CD146 negative subset of marrow fibroblasts involved with cytokine production. *PLoS One* 2014;9:e109304.
45 Behrendt P, Häfelein K, Preusse-Prange A, *et al*. Il-10 ameliorates TNF-α induced meniscus degeneration in mature meniscal tissue in vitro. *BMC Musculoskelet Disord* 2017;18:197.
46 Lee CH, Rodeo SA, Fortier LA, *et al*. Protein-releasing polymeric scaffolds induce fibrochondrocytic differentiation of endogenous cells for knee meniscus regeneration in sheep. *Sci Transl Med* 2014;6:266ra171.
47 Zhang Z-Z, Chen Y-R, Wang S-J, *et al*. Orchestrated biomechanical, structural, and biochemical stimuli for engineering anisotropic meniscus. *Sci Transl Med* 2019;11:eaao0750.
48 MacBarb RF, Makris EA, Hu JC, *et al*. A chondroitinase-ABC and TGF-β1 treatment regimen for enhancing the mechanical properties of tissue-engineered fibrocartilage. *Acta Biomater* 2013;9:4626–34.

CLINICAL SCIENCE

Development of a prediction model for inpatient gout flares in people with comorbid gout

Kanon Jatuworapruk ,[1,2] Rebecca Grainger,[1] Nicola Dalbeth,[3] William J. Taylor[1]

Handling editor Josef S Smolen

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-216277).

[1]Department of Medicine, University of Otago, Wellington, New Zealand
[2]Department of Medicine, Faculty of Medicine, Thammasat University, Pathumthani, Thailand
[3]Department of Medicine, University of Auckland, Auckland, New Zealand

Correspondence to
Dr Kanon Jatuworapruk, Department of Medicine, University of Otago, Wellington, New Zealand; kanon@tu.ac.th

Received 6 September 2019
Revised 21 November 2019
Accepted 25 November 2019
Published Online First 6 December 2019

To cite: Jatuworapruk K, Grainger R, Dalbeth N, *et al*. *Ann Rheum Dis* 2019;**79**:418–423.

ABSTRACT

Objectives Hospitalisation is a risk factor for flares in people with gout. However, the predictors of inpatient gout flare are not well understood. The aim of this study was to develop a prediction model for inpatient gout flare among people with comorbid gout.

Methods We used data from a retrospective cohort of hospitalised patients with comorbid gout from Wellington, Aotearoa/New Zealand, in 2017 calendar year. For the development of a prediction model, we took three approaches: (A) a clinical knowledge-driven model, (B) a statistics-driven model and (C) a decision tree model. The final model was chosen based on practicality and performance, then validated using bootstrap procedure.

Results The cohort consisted of 625 hospitalised patients with comorbid gout, 87 of whom experienced inpatient gout flare. Model A yielded 9 predictors of inpatient gout flare, while model B and C produced 15 and 5, respectively. Model A was chosen for its simplicity and superior C-statistics (0.82) and calibration slope (0.93). The final nine-item set of predictors were pre-admission urate >0.36 mmol/L, tophus, no pre-admission urate-lowering therapy (ULT), no pre-admission gout prophylaxis, acute kidney injury, surgery, initiation or increase of gout prophylaxis, adjustment of ULT and diuretics prior to flare. Bootstrap validation of the final model showed adequate C-statistics and calibration slope (0.80 and 0.78, respectively).

Conclusion We propose a set of nine predictors of inpatient flare for people with comorbid gout. The predictors are simple, practical and are supported by existing clinical knowledge.

Key messages

What is already known about this subject?
► Hospitalisation increases the risk of gout flare but the factors associated with this risk are not well understood.

What does this study add?
► This study proposes a prediction model for flares in hospitalised patients with comorbid gout.

How might this impact on clinical practice or future developments?
► The study might assist clinicians to identify patients with high risk of inpatient gout flare.

INTRODUCTION

Gout is one of the most common inflammatory joint diseases.[1] In Aotearoa/New Zealand, there is an especially high prevalence among Māori and Pacific peoples.[2] Long-standing hyperuricemia plays a key role in the development of gout, leading to monosodium urate crystal deposition and subsequent acute inflammatory response (gout flare). People with comorbid gout are hospitalised more frequently than people who do not have gout.[3] Inpatient gout flare adds 3–6 days to an admission,[4 5] and increases healthcare costs.[6]

Many factors are believed to be associated with inpatient gout flare, including urate-lowering therapy (ULT) withdrawal, exposure to diuretics, overhydration, acidaemia and surgery.[7] However, evidence linking these potential factors to the occurrence of inpatient gout flare is limited to a few studies conducted among different subsets of hospitalised gout patients.

A Korean study compared 67 people with post-surgical gout flare with 67 people without post-surgical flare. Three-day pre-surgical serum urate ≥0.54 mmol/L and cancer surgery were associated with gout flare, with OR and 95% CI of 8.2 (2.2 to 30.5) and 6.2 (1.9 to 19.9), respectively.[8] Another study conducted in acute stroke patients compared 60 people with gout flare with 860 people without flare. The study found that history of gout (OR 14.3 (95% CI 6.75 to 30.18)), higher inpatient serum urate (OR 1.5 (95% CI 1.26 to 1.78)) and hypercholesterolaemia (OR 2.0 (95% CI 1.06 to 3.83)) were associated with inpatient gout flare.[4] A recent study using community-based gout cohort from the USA evaluated multiple potential predictors, including ULT use, ULT withdrawal, diuretics and serum urate at gout diagnosis, and did not find any association between these predictors and inpatient gout flare. However, the study recorded only 23 flare episodes, which may have been underpowered to detect a statistically significant association.[9]

Given the sparse evidence regarding the factors associated with inpatient gout flare, we undertook this study to identify the predictors of inpatient gout flare, and develop a prediction model for inpatient gout flare among people with comorbid gout.

METHODS

Study design and population

We used data from a retrospective cohort of hospitalised patients with comorbid gout from Wellington, Aotearoa/New Zealand. The population of interest was patients aged 18 years or older with comorbid gout discharged from hospitals in

the Wellington Region (Wellington, Kenepuru and Hutt hospitals) in 2017 calendar year. We defined comorbid gout as having received the diagnosis of 'gout', 'gouty arthritis', 'chronic gout' or 'tophaceous gout' as comorbid disease in the hospital records or discharge letter. For patients with more than one admission in the study period, only data from the first admission were analysed. Exclusion criteria included those having gout or gout-related complications as the primary admission diagnosis or receiving a gout diagnosis for the first time. The primary outcome was the development of inpatient gout flare. Gout flare was defined as a new episode of joint pain and swelling judged to be gout by the attending doctors or consultant rheumatologist or the episode satisfied the 2015 ACR/EULAR gout classification criteria.[10] When there were more than one flare episodes in an admission, only the first was analysed. Patients were divided into a 'flare group' and 'non-flare group' for analysis.

Patient identification and data collection

We identified potentially eligible patients using two methods. First, we identified all admissions that received the discharge comorbid diagnosis of M10 (Gout) according to the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification (ICD-10-AM). Comorbid diagnoses are recorded separately from the primary admission diagnosis, but tend to be recorded only when the comorbidity influenced the inpatient stay in some way. The second method was the 'word search', using structured query language to search for specific words which could appear anywhere within the electronic discharge letters. Search terms included 'gout', 'allopurinol', 'febuxostat' and 'colchicine'. The search yielded a list of admissions with discharge letters containing at least one of the words of interest. Data were manually collected from physical hospital records and the electronic laboratory database, which covers all clinical laboratory tests in the Wellington region. We reviewed hospital records in alphabetical order of the patient's unique National Health Index number until we reached the target sample size.

Sample size

We originally considered events per variables (EPV) ratio between 5 and 10 acceptable, with EPV of 10 as the optimal number to minimise overfitting of the regression model.[11] According to this rule, we needed 50–100 inpatient flare episodes to evaluate 10 candidate predictors. Assuming that the prevalence of inpatient flare was 35% among patients with comorbid gout,[12] a total sample size of at least 300 would suffice. To ensure adequate number of events, we decided to collect data of 600 individuals.

Variables

Fifty-two variables were collected (table 1). Comorbidities were based on the 18-item Functional Comorbidity Index[13] and other conditions related to gout. The primary admission diagnoses were grouped by organ system according to ICD-10 coding. If flare occurred, we recorded only data available prior to the flare episode. Drug adjustment was defined as any change of the dosage made before flare, except for gout prophylaxis which was categorised as started/increased, stopped/decreased or no adjustment. For the serum urate level, we reviewed the laboratory database for results within 12 months prior to admission. We recorded the variable as highest pre-admission urate >0.36 mmol/L, ≤0.36 mmol/L or not tested. We chose 0.36 mmol/L as cut point in accordance with the Study for Updated Gout Classification Criteria.[14]

Missing data

Missing data occurred only in the ethnicity variable; 20 of 625 patients (3%) included in the final analysis did not have ethnicity recorded. However, this did not affect our analysis, as we classified these patients as non-Māori/non-Pacific.

Model development

To ensure the robustness and validity of the prediction model, we took three different approaches of model development: a clinical knowledge-driven model (model A), a statistics-driven model (model B) and a decision tree model (model C). To correct

Table 1 Candidate variables*

Domains	Variables
Demographics	(1) age ≥65 years, (2) male, (3) Māori or Pacific ethnicity
Comorbidities	(4) AKI (receiving diagnosis of 'AKI' or 'acute renal failure'),† (5) urinary tract stone, (6) hypercholesterolaemia (cholesterol level >5.2 mmol/L), (7) alcohol drinking (yes/no)† FCI: (8) arthritis (rheumatoid or osteoarthritis), (9) osteoporosis, (10) asthma, (11) COPD, ARDS, emphysema, (12) angina, (13) CHF or heart disease, (14) heart attack (myocardial infarction), (15) neurological disease (multiple sclerosis or Parkinson's), (16) stroke or TIA, (17) PVD, (18) DM type I or II, (19) upper GI disease (ulcer, hernia, reflux), (20) depression, (21) anxiety or panic, (22) visual impairment (cataracts, glaucoma, macular degeneration), (23) hearing impairment (very hard of hearing, even with hearing aids), (24) DDD (back disease, spinal stenosis or severe chronic back pain), (25) obesity or body mass index >30, (26) FCI score
Admission	Primary admission diagnosis: (27) infectious disease, (28) neoplasm, (29) diseases of the blood, (30) diseases of nervous system, (31) diseases of circulatory system, (32) diseases of respiratory system, (33) diseases of digestive systems, (34) diseases of the skin, (35) diseases of musculoskeletal system, (36) diseases of genitourinary system, (37) injury/external causes Treatments (prior to flare): (38) diuretics adjustment during admission,† (39) warfarin adjustment during admission, (40) dialysis,† (41) surgery,† (42) IV fluid ≥2 L within first 48 hours,† (43) blood product within first 48 hours
Gout history	(44) tophus,† (45) no pre-admission ULT,† (46) no pre-admission prophylaxis use,† (47) in-admission ULT adjustment (prior to flare),† (48) in-admission gout prophylaxis adjustment (prior to flare)†
Laboratory	(49) pre-admission urate >0.36 mmol/L† Inflammatory markers prior to flare: (50) CRP ≥100 mg/L, (51) neutrophil ≥15×10^9/L, (52) platelet ≥450×10^9/L

*All variables were included in statistics-driven and decision tree model.
†Factors selected for the clinical knowledge-driven model.
AKI, acute kidney injury; ARDS, acquired respiratory distress syndrome; CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; CRP, C-reactive protein; DDD, degenerative disc disease; DM, diabetes mellitus;FCI, Functional Comorbidity Index; GI, gastrointestinal; IV, intravenous; PVD, peripheral vascular disease; TIA, transient ischaemic attack; ULT, urate-lowering therapy.

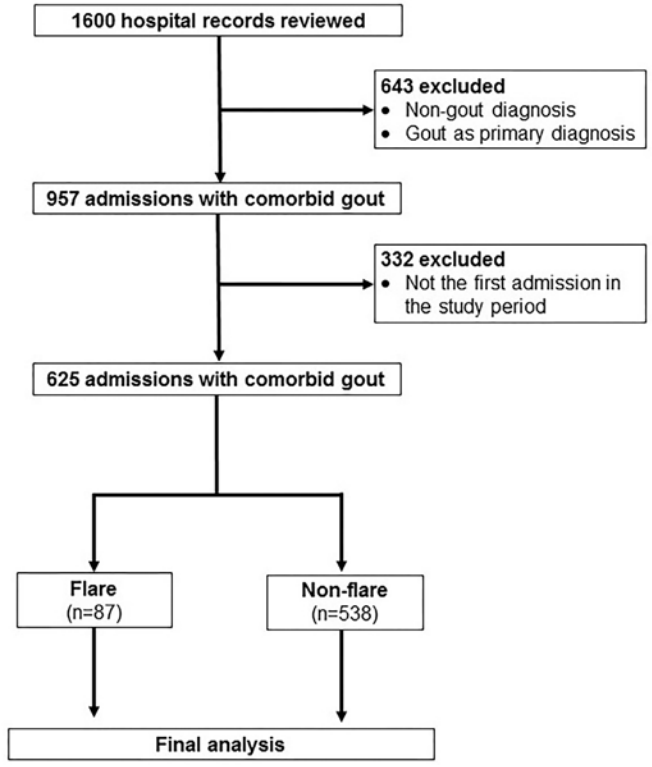


Figure 1 Study flow diagram.

for over-optimism, we applied the linear shrinkage technique to model A and the penalised regression model B.

Model A: clinical knowledge-driven model. We reduced the candidate variables by preselection based on (1) existing literature or well-established link to the fluctuation of serum urate level, which could theoretically cause gout flare and (2) their availability in routine hospital setting without requiring additional intervention to assess them. Knowledge-based variable selection ensured that all variables in the final model would make sense to clinicians. We then simultaneously entered all preselected variables into a standard logistic regression model, with inpatient flare (yes/no) as the dependent variable. The final set of variables included only those with p value<0.05 from the regression analysis. To correct for over-optimism, we further shrunk the regression coefficient by multiplication with a linear shrinkage factor. The shrinkage factor was derived from Van Houwelingen and Le Cassie's heuristic formula: s = [model χ^2 − (df − 1)]/model χ^2. Model χ^2 indicated χ^2 value of the model calculated from log-likelihood scale and df indicated degree of freedom.[15] We reported the shrunken coefficients (βs) for each variables in the final model.

Model B: statistics-driven model. This model relied solely on the statistical process to select the variables and estimate the regression coefficients. This was also an exploratory approach to identify potentially new predictors not previously proposed. We chose the least absolute shrinkage and selection operator (LASSO) procedure due to its ability in both variable selection and shrinkage. For a small data set with low EPV ratio, LASSO is preferable to the traditional stepwise regression, which has limited power and is prone to unstable selection.[16] LASSO is a type of penalised regression procedure that combines the estimation of regression coefficients with shrinkage and variable selection. The LASSO procedure corrects for over-optimism by constraining the regression coefficients using penalty factor (λ), derived from 10-fold cross-validation. In LASSO, the regression coefficients of some variables are shrunk to zero, effectively excluding them from the model.[17] We entered all 52 candidate variables into the LASSO model and set inpatient gout flare (yes/no) as the dependent variable.

Model C: decision tree model. A χ^2 automatic interaction detection (CHAID) decision tree is a type of supervised machine learning algorithm, which extracts a model from the observation of the cohort. A decision tree is non-parametric, meaning that it does not make prior assumptions about the data distribution.[18] The CHAID algorithm starts with a single node representing the entire cohort and splits the node using χ^2 test. Splitting continues on each successive node, creating a tree-like structure. The branching stops when the following stopping criteria applies: (1) no more than three levels of branches, (2) parent node size of at least 20 subjects and (3) child node size of at least 10 subjects. The minimal size of the node was determined by a rule of thumb that a node should not be smaller than 1% of the cohort.[19] We applied the algorithm to all 52 candidate variables, with inpatient gout flare (yes/no) as the dependent variable. We present the results as a decision tree, with each node representing corresponding variable category and the percentage with flare.

Table 2 Cohort demographics and admission data

Variables	Overall, n=625	Flare group, n=87	Non-flare group, n=538
Demographics			
Male, n (%)	487 (77.9)	75 (86.2)*	412 (76.6)
Age, mean±SD, year	68.9±13.6	69.3±13.2	68.8±13.7
Ethnicity (n=605)			
New Zealand European, n (%)	361 (59.7)	44 (51.8)	317 (61.0)
Māori, n (%)	115 (19.0)	16 (18.8)	99 (19.0)
Pacific, n (%)	103 (17.0)	18 (21.2)	85 (16.3)
Asian, n (%)	26 (4.3)	7 (8.2)	19 (3.7)
Admission			
Length of stay, median (IQR), day	3 (6)	8 (13)†	2 (4)
In-hospital mortality, n (%)	10 (1.6)	1 (1.1)	9 (1.7)
Most common categories of primary admission diagnosis			
Circulatory system, n (%)	168 (26.8)	21 (24.2)	147 (27.4)
Digestive system, n (%)	78 (12.5)	11 (12.6)	67 (12.5)
Respiratory system, n (%)	75 (12.0)	14 (16.2)	61 (11.3)
Musculoskeletal system, n (%)	65 (10.4)	6 (6.9)	59 (11.0)
Genitourinary system, n (%)	51 (8.2)	9 (10.3)	42 (7.8)

*P <0.05, compared with non-flare group.
†P<0.001, compared with non-flare group.
IQR, interquartile range.

Model performance and selection

We tested for the model's discrimination and calibration using C-statistics and calibration slope, respectively. The C-statistic is the area under the receiver operating characteristic curve created by plotting the true positive rate against the false positive rate of the model. C-statistics range between 0.5 and 1.0, where the latter indicates perfect discrimination between flare and non-flare group. The calibration slope examines the relationship between the predicted probability (x-axis) and the observed probability of flare (y-axis). We grouped the patients by deciles of their predicted probability of flare and plotted the average values of each group against the group's observed probability.

Table 3 The three approaches for model development and their performance

Approaches	Variables (n)	Variables	C-statistics (95% CI)	Calibration slope (95% CI)
Model A: clinical knowledge-driven	9	No pre-admission ULT,* ULT adjustment,*† diuretic adjustment,*† pre-admission urate>0.36 mmol/L,* tophus, no pre-admission prophylaxis, prophylaxis started/increased,† AKI, surgery, IV fluid≥2 L within 48 hours,‡ dialysis,‡ alcohol drinking (yes/no)‡	0.82 (0.77 to 0.86)	0.93 (0.33 to 1.52)
Model B: statistics-driven	15	No pre-admission ULT,* ULT adjustment,*† diuretic adjustment,*† pre-admission urate>0.36 mmol/L,* tophus, no pre-admission prophylaxis, prophylaxis started/increased,† AKI, IV fluid≥2 L within 48 hours, disease of the nervous system (primary diagnosis), stroke (comorbidity), warfarin adjustment, blood product, CRP>100 mg/L† or CRP not tested, platelet>450×10^9/L†	0.81 (0.76 to 0.86)	2.11 (1.66 to 2.56)
Model C: decision tree	5	No pre-admission ULT,* ULT adjustment,*† diuretic adjustment,*† pre-admission urate>0.36 mmol/L,* IV fluid≥2 L within 48 hours	0.76 (0.71 to 0.82)	NA

*Selected by all three approaches.
†Occurs during admission and prior to flare.
‡Removed from the multivariate regression model due to p-value greater than 0.05.
AKI, acute kidney injury;CRP, C-reactive protein; IV, intravenous; NA, not applicable; ULT, urate-lowering therapy.

A calibration slope of 1.0 indicates perfect agreement between the predicted and the observed probability of flare. Assessment of calibration of model C was not applicable, as a decision tree does not produce regression coefficients. We planned to select the model with the best discrimination, calibration and practicality. We also planned to select a model which is simple enough for healthcare professionals to use in clinical practice without requiring additional interventions (eg, no additional blood tests required).

Model validation

We used a bootstrap procedure for model validation. The bootstrap procedure produces large number of similar but not identical versions of the original data sets by performing random resampling with replacements.[20] We generated 1000 bootstrap samples and derived the model from each sample using the same process performed for model development. We then applied each bootstrap model to the original data set and calculated the bootstrap C-statistics and calibration slope. The bootstrap procedure estimated the average optimism across the bootstrap samples and produced the optimism-corrected calibration slope and C-statistics. An ideal model would have the optimism-corrected performances similar to the original values, indicating that the model performed similarly in different hypothetical data sets.

Statistical analysis was performed using RStudio (V.1.2.1335) and IBM SPSS Statistics software (V.25).

Patient and public involvement

It was not appropriate or possible to involve patients or the public in the design, conduct, reporting or dissemination of our research.

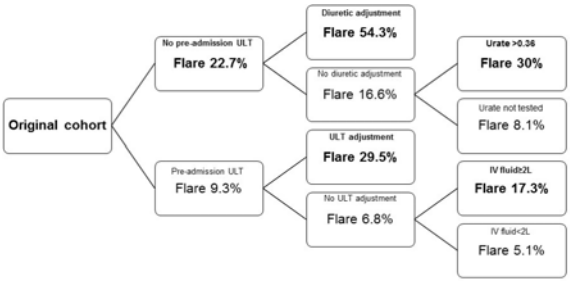


Figure 2 χ^2 automatic interaction detection decision tree.

RESULTS

Cohort

Data of 625 patients were included in the final analysis. There were 87 (14%) inpatient gout flare episodes. The majority of the cohort were men (78%). Hospital length of stay was significantly longer in the flare group compared with non-flare group (median 8 vs 2 days, p<0.001). Study flow diagram and characteristics of the cohort are shown in figure 1, table 2 and online supplementary table 1.

Prediction models

The three approaches resulted in three different sets of variables associated with inpatient gout flare (table 3). Four variables were consistently selected across all three models: 'no pre-admission ULT', 'ULT adjustment ', 'diuretics adjustment ' and 'pre-admission urate>0.36 mmol/L'. In the clinical knowledge-driven approach (model A), we included a preselected set of 12 candidate variables (table 1), with the results indicating that 9 were associated with inpatient gout flare. The statistics-driven approach (model B) produced 15 variables. The regression coefficients for each variables in models A and B are shown in online supplementary table 2. The decision tree approach (model C) produced a decision tree which described the relationship between five variables and the risk of inpatient gout flare (figure 2).

Models A and B showed comparable discrimination, with C-statistics 0.82 (95% CI 0.77 to 0.86) and 0.81 (95% CI 0.76 to 0.86), respectively. Model C had slightly inferior discrimination at 0.76 (95% CI 0.71 to 0.82). However, model A had the most optimal model fit, with calibration slope of 0.93 (95% CI 0.33 to 1.52) and the predicted probabilities of flare between the lower and upper deciles of the calibration plot were 1% and 48%, respectively (online supplementary figure 1). Considering practicality, models A and C were simpler than model B, containing fewer predictors which were readily available in hospital setting and did not require additional blood tests. Ultimately, we selected model A for its superior performance and simplicity. Table 4 shows the details of the final model.

Finally, we performed a 1000-sample bootstrap validation of the selected model A. The optimism-corrected C-statistics was 0.80 (95% CI 0.78 to 0.88), which was comparable with the original value of 0.82. The optimism-corrected calibration slope was 0.78 (95% CI 0.52 to 1.02), compared with the original slope of 0.93. In other words, the final model showed similar

Table 4 Variables in the final model using multivariable regression with shrinkage

Variable	OR (95% CI)	P value	Regression coefficients	Regression coefficients with shrinkage*
Intercept	–	–	−6.060	−5.393
No pre-admission ULT	4.40 (2.50 to 7.87)	<0.001	1.481	1.318
ULT adjustment†	3.04 (1.31 to 7.03)	0.009	1.111	0.989
Diuretics adjustment†	2.91 (1.58 to 5.39)	0.001	1.070	0.952
Pre-admission urate>0.36 mmol/L	3.36 (1.31 to 8.61)	0.012	1.212	1.079
Tophus	4.32 (1.39 to 13.40)	0.011	1.462	1.301
No pre-admission gout prophylaxis	8.44 (2.26 to 31.57)	0.002	2.133	1.898
Gout prophylaxis started or increased†	17.36 (2.76 to 109.24)	0.002	2.854	2.540
Acute kidney injury	2.33 (1.23 to 4.43)	0.01	0.845	0.752
Surgery	1.84 (1.01 to 3.38)	0.049	0.610	0.543

*Shrinkage factor=0.89.
†Occurs during admission and prior to flare.
ULT, urate-lowering agent.

discrimination but slightly lower calibration when tested in 1000 hypothetical data sets.

DISCUSSION

We found that nine variables from the clinical knowledge-driven model were associated with the risk of inpatient gout flare among patients with comorbid gout. These variables were consistent with existing knowledge. In theory, an acute illness could prime macrophages for an inflammatory response to deposited monosodium urate crystals.[21] In addition, a change in serum urate concentration could also promote partial dissolution of crystal deposits and crystal release.[7] In our study, the association between gout flare and ULT adjustment, diuretic adjustment, surgery and AKI may be explained by these mechanisms, as these factors were likely associated with the patients' acute illness and alteration of serum urate level. The effect of ULT adjustment on gout flare is further supported by trials that found an increase of gout flare rate after initiation of ULT.[22 23] The remaining five predictors (no pre-admission ULT or gout prophylaxis, gout prophylaxis started/increased during admission, tophus and urate>0.36 mmol/L) generally reflect suboptimal gout treatment. Prophylaxis started/increased during admission, in particular, likely reflects poorly controlled gout (ie, frequent flare) that prompted the doctors to take preventive action. In the context of multivariable model, the presence of 'prophylaxis started/increased' further emphasises the significance of the other eight predictors because they remain predictive of flare despite some people receiving new or higher dose of prophylaxis. Urate concentration of <0.36 mmol/L is recommended as a therapeutic target of gout management.[24 25] Our study strengthens the importance of this treat-to-target strategy.

This study took many steps to minimise potential bias. We identified patients using word search and ICD-10 coding for comorbidity. Patient identification using discharge ICD coding alone could carry bias towards flare group, as gout is more likely to be coded if it was active during inpatient stay. Previous studies relying on ICD coding alone reported much higher prevalence of inpatient flare among people with comorbid gout, compared with our cohort (34%–35% vs 14%).[12 26] We also excluded repeated admissions to ensure that all cases in the final cohort were independent from each other. Finally, all comorbid gout and flare episodes were confirmed by manual hospital record review in accordance to the predefined case definition.

To minimise over-optimism, we applied shrinkage techniques to correct for potential over-optimism in models A and B. We refrained from splitting our cohort into 'development' and 'validation' subgroups, as data splitting could undermine the power of regression analysis and might be prone to fortuitous splitting. For a data set with relatively small event number, bootstrap validation is a preferable option for internal validation.[20] Furthermore, our decision to take three different approaches of model development allowed us to see patterns of variables emerging from different methods. For the nine-item final model, four variables were also selected by the other two approaches. This apparent agreement between different methods is relatively reassuring in terms of the reliability of the models.

The final model was derived from a clinical knowledge-driven variable selection. This approach is generally preferable to selection based on p value alone (univariate and stepwise selection), which is prone to selecting spurious predictors and overfitting.[16] In this regard, the clinical knowledge-driven model could be considered as more robust than the statistics-driven model. This assumption is supported by model A's superior performance. Another advantage of the clinical knowledge-driven approach is that it ensures that all variables are intuitive to clinicians and that they are feasible in routine hospital setting.

There are a number of limitations to our study. It must be emphasised that a predictive model cannot capture all existing predictors. The retrospective nature of this study meant that data were limited to those documented in hospital record. Some variables, such as tophus and alcohol history, might be underestimated or otherwise inaccurate especially in the non-flare group where gout was not the focus during admission. There was a small possibility that people with conditions which often mimic the natural course of gout flare (eg, calcium pyrophosphate crystal arthritis) could have been mistakenly included as having a gout flare episode. However, this could be considered a strength. If the cohort were confounded by non-gout episodes, the association of the selected covariates should have been weaker because our covariates are highly specific to gout. In our opinion, the strong association found in our models despite this limitation is compelling. Regarding sample size, our cohort was relatively small for prediction model development. Despite our effort to make the model as robust as possible, there remains a probability that the model will not perform as well as expected in a different population. The number of gout flare episodes were also lower than originally expected, which could limit the power of the regression analysis. An external validation is planned for a future study.

Our study provides some evidence supporting the role of several variables long suspected to increase the risk of inpatient gout flare. The proposed predictors are practical and

non-invasive, as they are readily available in typical hospital care setting. The model may help clinicians identify hospitalised patients with comorbid gout who are at risk of developing flare. It must be emphasised that the target group for these predictors is hospitalised people with comorbid gout rather than the general inpatient population. Consequently, identification of patients with comorbid gout at the time of admission is essential before any risk evaluation could begin.

Acknowledgements We would like to thank Gordon Purdie and James Stanley, Biostatistical Services, University of Otago, Wellington, for their statistical advices.

Contributors KJ and WT designed the study. KJ collected the data. KJ, RG, ND and WT were involved in the data analysis, data interpretation and model development. All authors were involved in manuscript preparation and have approved the submitted version of the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests RG reports grants and personal fees from AbbVie, personal fees from Janssen and personal fees from Pfizer, outside submitted work. ND reports grants and personal fees from AstraZeneca, grants from Amgen, personal fees from Dyve, personal fees from Hengrui, personal fees from Horizon, personal fees from Kowa, personal fees from Abbvie, personal fees from Pfizer, personal fees from Janssen, outside the submitted work. KJ and WT have nothing to disclose.

Patient consent for publication Not required.

Ethics approval The Human Research Ethics Committee, University of Otago, reviewed and approved the study protocol (reference number H18/012).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iD

Kanon Jatuworapruk http://orcid.org/0000-0002-9301-8941

REFERENCES

1 Kuo C-F, Grainge MJ, Zhang W, *et al*. Global epidemiology of gout: prevalence, incidence and risk factors. *Nat Rev Rheumatol* 2015;11:649–62.
2 Winnard D, Wright C, Taylor WJ, *et al*. National prevalence of gout derived from administrative health data in Aotearoa New Zealand. *Rheumatology* 2012;51:901–9.
3 Singh JA, Strand V. Gout is associated with more comorbidities, poorer health-related quality of life and higher healthcare utilisation in US veterans. *Ann Rheum Dis* 2008;67:1310–6.
4 Lin Y-H, Hsu H-L, Huang Y-C, *et al*. Gouty arthritis in acute cerebrovascular disease. *Cerebrovasc Dis* 2009;28:391–6.
5 Lee G, Roberts L. Healthcare burden of in-hospital gout. *Intern Med J* 2012;42:1261–3.
6 Saseen JJ, Agashivala N, Allen RR, *et al*. Comparison of patient characteristics and gout-related health-care resource utilization and costs in patients with frequent versus infrequent gouty arthritis attacks. *Rheumatology* 2012;51:2004–12.
7 Fisher MC, Pillinger MH, Keenan RT. Inpatient gout: a review. *Curr Rheumatol Rep* 2014;16:458.
8 Kang EH, Lee EY, Lee YJ, *et al*. Clinical features and risk factors of postsurgical gout. *Ann Rheum Dis* 2008;67:1271–5.
9 Zleik N, Elfishawi MM, Kvrgic Z, *et al*. Hospitalization increases the risk of acute arthritic flares in gout: a population-based study over 2 decades. *J Rheumatol* 2018;45:1188–91.
10 Neogi T, Jansen TLTA, Dalbeth N, *et al*. 2015 gout classification criteria: an American College of Rheumatology/European League against rheumatism collaborative initiative. *Arthritis Rheumatol* 2015;67:2557–68.
11 Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and COX regression. *Am J Epidemiol* 2007;165:710–8.
12 Petersel D, Schlesinger N. Treatment of acute gout in hospitalized patients. *J Rheumatol* 2007;34:1566–8.
13 Groll DL, To T, Bombardier C, *et al*. The development of a comorbidity index with physical function as the outcome. *J Clin Epidemiol* 2005;58:595–602.
14 Taylor WJ, Fransen J, Jansen TL, *et al*. Study for updated gout classification criteria (sugar): identification of features to classify gout. *Arthritis Care Res* 2016;68:1894–8.
15 Van Houwelingen JC, Le Cessie S. Predictive value of statistical models. *Stat Med* 1990;9:1303–25.
16 Steyerberg EW, Eijkemans MJ, Harrell FE, *et al*. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. *Stat Med* 2000;19:1059–79.
17 Steyerberg EW, Eijkemans MJC, Habbema JDF. Application of shrinkage techniques in logistic regression analysis: a case study. *Stat Neerl* 2001;55:76–88.
18 Geurts P, Irrthum A, Wehenkel L. Supervised learning with decision tree-based methods in computational and systems biology. *Mol Biosyst* 2009;5:1593–605.
19 Song Y-Y, Lu Y. Decision tree methods: applications for classification and prediction. *Shanghai Arch Psychiatry* 2015;27:130–5.
20 Moons KGM, Kengne AP, Woodward M, *et al*. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart* 2012;98:683–90.
21 So AK, Martinon F. Inflammation in gout: mechanisms and therapeutic targets. *Nat Rev Rheumatol* 2017;13:639–47.
22 Schumacher HR, Becker MA, Wortmann RL, *et al*. Effects of febuxostat versus allopurinol and placebo in reducing serum urate in subjects with hyperuricemia and gout: a 28-week, phase III, randomized, double-blind, parallel-group trial. *Arthritis Rheum* 2008;59:1540–8.
23 Becker MA, Schumacher HR, Espinoza LR, *et al*. The urate-lowering efficacy and safety of febuxostat in the treatment of the hyperuricemia of gout: the confirms trial. *Arthritis Res Ther* 2010;12.
24 Khanna D, Fitzgerald JD, Khanna PP, *et al*. 2012 American College of rheumatology guidelines for management of gout. Part 1: systematic nonpharmacologic and pharmacologic therapeutic approaches to hyperuricemia. *Arthritis Care Res* 2012;64:1431–46.
25 Richette P, Doherty M, Pascual E, *et al*. 2016 updated EULAR evidence-based recommendations for the management of gout. *Ann Rheum Dis* 2017;76:29–42.
26 Gnanenthiran SR, Hassett GM, Gibson KA, *et al*. Acute gout management during hospitalization: a need for a protocol. *Intern Med J* 2011;41:610–7.

LACC1 gene mutation in three sisters with polyarthritis without systemic features

Juvenile idiopathic arthritis (JIA) refers to a group of disorders characterised by wide phenotypic diversity and genetic heterogeneity. Disordered immune response to an environmental trigger in a genetically predisposed individual is the proposed mechanism for most JIA subtypes.[1 2] There are emerging reports on new gene locus being identified especially in families with many affected members.[3 4] We report three sisters with polyarthritis who were identified to have causative variant in Laccase domain containing one (*LACC1*) gene by whole exome sequencing.

CASE DETAILS

A non-consanguineous family from north-western part of India reported with three daughters (P1, P2 and P3) having polyarticular joint disease. The onset of symptoms in the third child (11 months) prompted the parents to seek medical help. The eldest sister was 5 years and 9 months old, while the second daughter was 3 years old at the time of presentation. The chronology of symptoms was similar in all three children. Joint symptoms started in infancy at around 9–10 months of age with involvement of knee and ankle and rapidly progressed to involve small joints and cervical spine too (figure 1A–C). Over the next 2 years, multiple joint involvement, pain, deformities and contractures resulted in limitation of normal routine activities, and the children were bed bound. There was no history suggestive of psoriasis or inflammatory bowel disease in family. On examination they were stunted, no facial dysmorphism was noted, had nail dystrophy and marked swelling and deformity of large and small joints. The investigations at the time of first evaluation are tabulated in the online supplementary table 1. Disease activity based on Juvenile Arthritis Disease Activity Score (JADAS-27) was 49.4, 46.6 and 45 in P1, P2 and P3, respectively (score range 0-57). Radiographs in all three sisters showed osteopenia and erosion of vertebrae without any platyspondyly (see online supplementary figure 1).

Initial clinical possibilities considered were Torg Winchester syndrome, Pseudorheumatoid chondrodysplasia and Familial inflammatory arthropathy. However, they fulfilled the International League of Associations for Rheumatology classification for polyarticular JIA.[5] Non-steroidal anti-inflammatory drugs (naproxen) was initiated initially, and they reported relief of pain and improvement in restriction. Later, in view of rheumatoid factor being positive in eldest sister, a trial of oral prednisolone (1 mg/kg) followed by subcutaneous methotrexate (10 mg/m^2/week) resulted in marked clinical improvement. Younger sisters were also treated on the same lines. All three have significantly improved and are currently ambulatory. Investigations at 2-year follow-up have been given in the online supplementary table 2. Disease activity based on JADAS-27 was 13, 7 and 7 in P1, P2 and P3, respectively. However, they are stunted with poor growth velocity.

Sanger sequencing analysis for *MMP2* (Torg Winchester syndrome) and *WISP3* (pseudorheumatoid chondrodysplasia) did not reveal any significant variant. Subsequently at 2-year follow-up, whole exome sequencing in P1 and P2 revealed a homozygous variant, c.832G>C; p.(Ala278Pro) in exon 4 of *LACC1* (NM_001128303.2). The same variant was also present in homozygous state in P3 and heterozygous state in both the parents, as confirmed by Sanger Sequencing (figure 1D). This variant has not been observed in population databases like 1000 Genomes Project, ExAC Browser and our in-house data of 592 exomes in homozygous state. Multiple in silico prediction tools (SIFT,[6] Mutation Taster[7] and Polyphen[8]) are consistent in predicting that this variant is damaging to LACC1 protein function. *LACC1* is involved in inflammasome activation, fatty acid oxidation and production of reactive oxygen species by nicotinamide adenine dinucleotide phosphate hydrogen (NADPH) oxidase and mitochondrial pathways.[9 10] Mutation in this gene has been linked to JIA, inflammatory bowel disease and Behcet's disease.[4 11 12]

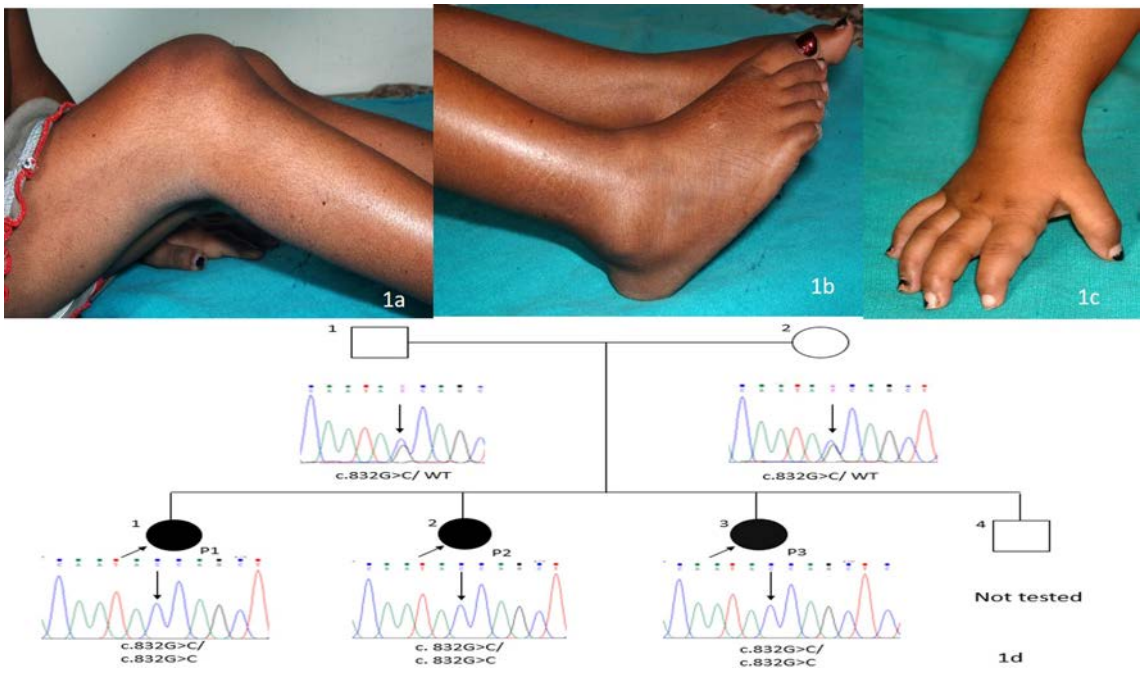


Figure 1 (A) Knee joint swelling in P1. (B) Swelling in ankle joint (P2). (C) Swelling in wrist and finger joints (P3). (D) Electropherogram of three siblings and parents. P1, patient 1; P2, patient 2; P3, patient 3.

Table 1 Summary of clinical and molecular features of patients with *LACC1* gene variants

Author/ year	Total number of patients (sex)	Mean age	Fever	Large joints	Small joints	Rash	Type of JIA	Inflammatory markers	Rheumatoid factor	*LACC1* gene variants
Wakil *et al* 2015[3]	13 (all females)	3.2±1.8 years	+	+	+	+	Polyarthritis sJIA	+	+	c.T850C p.C284R
Kallinich *et al* 2016[4]	2 (all females)	Patient 1: 15 months	–	+	–	–	Polyarthritis	+	–	c.827delC.
		Patient 2: 16 months	–	+	–	–	Polyarthritis	+	–	c.827delC.
Karacan *et al* 2018[13]	17 (10 males and 7 females)	15±9.1 years	–	+	+	–	Polyarthritis, oligoarthritis and ethesitis related arthritis	+	–	Family A- c.3G>A (p.0) FamilyB-c.1240C>T p.(Arg414Ter) Family C- c.988_990del p.(Ile330del) Family F- c.1109G>A p.(Cys370Tyr)]
Aroategui *et al* 2015[14]	3	3 years	–	+	+	–		+		c.128_129delGT (p.Cys43Tyrfs*6)
Index patients	3 (all females)	11±4 years	–	+	+	–	Polyarthritis	+	+	c.832G>C; p.(Ala278Pro)

JIA, juvenile idiopathic arthritis.

JIA is mostly sporadic and familial aggregation is uncommon. However, there have been few reported cases worldwide where familial patients of JIA were associated with variants in *LACC1*. These reports have been summarised in table 1.[3 4 13 14]

In our patients, all three sisters had polyarthritis onset in infancy and had similar pattern of joint involvement without fever, rash or serositis. They had raised inflammatory parameters and responded well to immunosuppressant, though complete remission has not been attained. The variant c.832G>C; p.(Ala278Pro) is present in the multicopperoxidase domain of *LACC1* which further supports its causative role.

It is interesting to note that patients with familial aggregation of JIA with *LACC1* mutation have shown varied pattern of involvement sJIA, polyarticular and extended oligoarticular patterns. However, all have elevated inflammatory parameters, thrombocytosis and response to immunosuppressants. The role of LACC1 protein in inflammasome formation could explain its association in various inflammatory disorders. However, the exact pathogenesis need to be studied.

CONCLUSION

This case report further supports the emerging evidence of causal role of pathogenic variants in *LACC1* with familial aggregation of JIA. Long-term follow-up of these patients may throw further highlight on the course of JIA in these subsets of patients.

Ankita Singh ,[1] **Deepti Suri,**[1] **Pandiarajan Vignesh,**[1] **Gummadi Anjani,**[2] **Prince Jacob,**[2] **Katta M Girisha**[2]

[1]Pediatric Allergy Immunology Unit, Advanced Pediatrics Centre, Post Graduate Institute of Medical Education and Research, Chandigarh, India
[2]Department of Medical Genetics, Kasturba Medical College, Manipal, Manipal Academy of Higher Education, Manipal, India

Correspondence to Dr Deepti Suri, Pediatric Allergy Immunology, Advanced Pediatrics Centre, Post Graduate Institute of Medical Education and Research, Chandigarh, 160012, India; surideepti@gmail.com

Correction notice This article has been corrected since it published Online First. The author affiliations have been updated.

Handling editor Josef S Smolen

Acknowledgements The authors are grateful to Department of Science and Technology, Government of India, Grant/Award Number: SB/SO/HS/005/2014 for funding the project.

Contributors All authors have contributed to the preparation of manuscript and in the patient management.

Funding Genetic analysis was done under the Department of Science and Technology, Government of India, Grant/Award Number: SB/SO/HS/005/2014.

Competing interests None declared.

Patient consent for publication Parental/guardian consent obtained.

Provenance and peer review Not commissioned; externally peer reviewed.

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-216263).

This work has previously been presented as a poster in EULAR 2019 congress and abstract has been published in Annals of Rheumatic diseases (FRI0578).

To cite Singh A, Suri D, Vignesh P, *et al*. *Ann Rheum Dis* 2020;**79**:425–426.

Received 10 September 2019
Revised 6 November 2019
Accepted 20 November 2019
Published Online First 6 December 2019

Ann Rheum Dis 2020;**79**:425–426. doi:10.1136/annrheumdis-2019-216263

ORCID iD
Ankita Singh http://orcid.org/0000-0001-7727-5553

REFERENCES

1 Moroldo MB, Tague BL, Shear ES, *et al*. Juvenile rheumatoid arthritis in affected sibpairs. *Arthritis Rheum* 1997;40:1962–6.
2 Prahalad S, Glass DN. A comprehensive review of the genetics of juvenile idiopathic arthritis. *Pediatric Rheumatology* 2008;6.
3 Wakil SM, Monies DM, Abouelhoda M, *et al*. Association of a Mutation in *LACC1* With a Monogenic Form of Systemic Juvenile Idiopathic Arthritis:Monogenic Systemic Juvenile Idiopathic Arthritis. *Arthritis & Rheumatology* 2015;67:288–95.
4 Kallinich T, Thorwarth A, von Stuckrad S-L, *et al*. Juvenile arthritis caused by a novel FAMIN (LACC1) mutation in two children with systemic and extended oligoarticular course. *Pediatr Rheumatol Online J* 2016;14.
5 Petty RE, Southwood TR, Manners P, *et al*. International League of associations for rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. *J Rheumatol* 2004;31:390–2.
6 Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 2009;4:1073–81.
7 Schwarz JM, Cooper DN, Schuelke M, *et al*. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods* 2014;11:361–2.
8 Adzhubei IA, Schmidt S, Peshkin L, *et al*. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248–9.
9 Sakurai T, Kataoka K. Basic and applied features of multicopper oxidases, CueO, bilirubin oxidase, and laccase. *Chem. Record* 2007;7:220–9.
10 Baldrian P. Fungal laccases – occurrence and properties. *FEMS Microbiol Rev* 2006;30:215–42.
11 Assadi G, Saleh R, Hadizadeh F, *et al*. LACC1 polymorphisms in inflammatory bowel disease and juvenile idiopathic arthritis. *Genes Immun* 2016;17:261–4.
12 Wu P, Du L, Hou S, *et al*. Association of *LACC1, CEBPB - PTPN1, RIPK2* and *ADO-EGR2* with ocular Behcet's disease in a Chinese Han population. *Br J Ophthalmol* 2018;102:1308–14.
13 Karacan I, Uğurlu S, Şahin S, *et al*. *LACC1* Gene Defects in Familial Form of Juvenile Arthritis. *J Rheumatol* 2018;45:726–8.
14 Arostegui JI, Rabionet R, Remesal A, *et al*. A family carrying a homozygous LACC1 truncated mutation expands the clinical phenotype of this disease beyond systemic-onset juvenile idiopathic arthritis. *Pediatric Rheumatology* 2015;13.

LRBA deficiency: a new genetic cause of monogenic lupus

Juvenile systemic lupus erythematosus (JSLE) is considered a polygenic disease, although identified causes of monogenic SLE and lupus-like syndrome are enlarging.[1 2] The genetic basis of polyautoimmune syndromes is also being elucidated, now including lipopolysaccharide-responsive beige-like anchor (LRBA) deficiencies.[3] We report a patient carrying a new deleterious LRBA mutation that associates with JSLE .

The patient, a girl born from healthy consanguineous parents, presented recurrent respiratory infections since 1 year of age, and since 7 years of age, chronic non-bloody diarrhoea diagnosed as non-specific colitis, IgA deficiency (<1.0 mg/dL) and arthralgia. At 10 years of age, laboratory tests revealed normal C1q, C2, C3, C4, CH50 and lymphocyte counts, homogeneous antinuclear antibody (ANA; HEp-2 1/640) and negative anti-dsDNA. Six months later, she presented respiratory distress, acute diarrhoea, pericardial effusion, serum glucose of 724 mg/dL, positive anti-IAA (42 IU/mL), and type 1 diabetes (T1D) was diagnosed. When 11 years of age, she clearly fulfilled the American College of Rheumatology 1997 SLE classification criteria by presenting polyarthritis, pericarditis, autoimmune haemolytic anaemia, alopecia, persistent malar rash, homogeneous ANA (1/1280), positive anti-dsDNA (119 IU/mL, ELISA, confirmed by *Crithidia luciliae*), anticardiolipin IgG, anti-thyroglobulin and anti-thyroid peroxidase. Prednisone and hydroxychloroquine were initiated. During the subsequent years, she presented a series of respiratory infections and died at 20 years of age due to pneumonia.

Whole exome sequencing (WES) from stored genomic DNA[4] identified a 1 bp deletion in *LRBA*, resulting in a frameshift with introduction of a premature stop codon (c.6742delT:p.Trp2248Glyfs*26; NM006726). The mutation appears novel as per databases (EXAC, GnomAD, 1000 Genomes, ESP6500) and literature search. The deletion was found in homozygosity in both index patient and her sister, who presented with recurrent bronchitis and pneumonias since the age of 2, bloodless diarrhoea episodes since 4 years of age, was diagnosed with Graves' disease at 10 years of age, suspected Beçhet's disease at 24 years and soon after died of respiratory complications. Both parents and two brothers were healthy and confirmed heterozygous carriers (figure 1A,B). The deletion is located in a highly conserved region within the Beige and Chediak-Higashi (BEACH) domain (figure 1C), classified as deleterious (combined annotation dependent depletion (CADD) score 39) and predicted in-silico to result in non-sense-mediated messenger RNA decay (using SIFTIndel and MutationTaster). Western blot analysis evidenced the LRBA protein was near absent in the homozygous sister but not in heterozygous relatives (figure 1D). Further WES analysis at genes associated with primary immunodeficiency, immune regulation, inflammatory bowel disease, T1D, IgA deficiency, monogenic lupus or SLE, revealed other mutations (see online supplementary table S1). These variants were all in heterozygosity excepted for a homozygous non-frameshift insertion in ATXN1, predicted in-silico to be neutral. ATXN1 has been associated with SLE and not with monogenic SLE. Overall, these additional mutations are

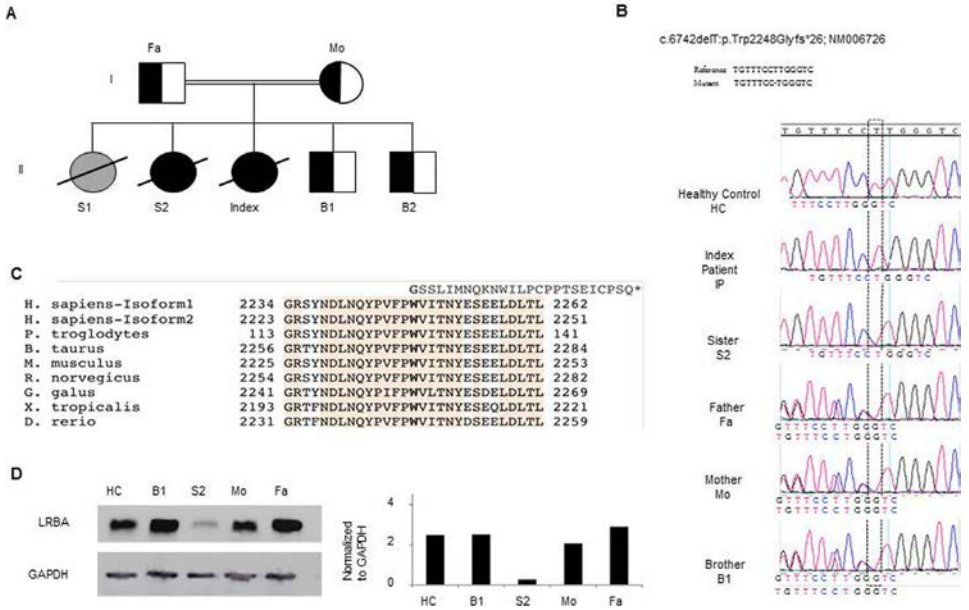


Figure 1 LRBA allele transmission analysis and protein production defect. (A) Pedigree of the index patient family indicating both LRBA genotype and health status. Double line, consanguineous; half-filled, heterozygous mutation and healthy; filled black, homozygous mutation and symptomatic; slashed symbol, deceased; filled grey, not genotyped (S1 died at 3 months of age with jaundice and unidentified liver disorder). (B) Electropherogram from Sanger sequencing of the LRBA exon 45 (shown is a selected area around the deletion) from the indicated family members and one unrelated healthy control. (C) Alignment of the LRBA protein BEACH domain from the indicated species. The amino acid sequence introduced by the frameshift mutation (*) is displayed above. Tryptophan (W) mutated in the reported patient is indicated in bold. Conserved amino acids are coloured. (D) Immunobloting of LRBA protein (~319 kDa) and GAPDH (~36 kDa) in PHA-stimulated PBMC lysates from the indicated family members and one unrelated healthy control (left) as well as densitometry quantification of LRBA protein normalised against GAPDH (right). BEACH, Beige and Chediak-Higashi; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; LRBA, lipopolysaccharide-responsive beige-like anchor; PBMC, peripheral blood mononuclear cell; PHA, phytohemagglutinin.

unlikely to be causal of the polyautoimmunity or the individual diseases presented, but some may have contributed to specify the latter.

The clinical features presented by the index patient and sister, excepted for JSLE, and the phenotypic variability between siblings, are strikingly similar to those reported for LRBA-deficient patients.[3 5] LRBA promotes CTLA4 expression, T cells from patients with SLE fail to upregulate CTLA4 on activation and GWAS associated CTLA4 with SLE.[6] These evidences suggest that alteration in CTLA4 pathway maybe causal for all clinical manifestations presented by the index patient, including JSLE. As both patients died during the course of this study, confirmation of this scenario will require a battery of assays following genome-editing techniques. In turn, these will inform whether CTLA4-Ig (abatacept) therapy may be considered for similar complex clinical presentations.

In conclusion, our study adds JSLE to the list of pathologies associated with LRBA deficiency and reinforces the notion that severe defects in immune regulation can lead to complex and multifaceted syndromes.

Bernadete L Liphaus ,[1] **Iris Caramalho,**[2] **Andreia Rangel-Santos,**[1] **Clóvis A Silva,**[1] **Jocelyne Demengeot,**[2] **Magda Maria Salles Carneiro-Sampaio**[3]

[1]Pediatrics, Children's Hospital, São Paulo, Brazil
[2]Instituto Gulbenkian de Ciência, Oeiras, Portugal
[3]Department of Pediatrics, Universidade de Sao Paulo, Sao Paulo, Brazil

Correspondence to Dr Bernadete L Liphaus, Pediatrics, Children's Hospital, São Paulo 05403-900, Brazil; bernadete.liphaus@hc.fm.usp.br

Handling editor Josef S Smolen

Acknowledgements Authors thank FAPESP for their financial support and Dr Adriana Almeida de Jesus for gathering patients with JSLE and associated immunodeficiencies.

Contributors BLL contributed to study conception, design and intellectual content, to data interpretation and wrote manuscript. IC analysed data and contributed to study intellectual content and manuscript writing. ARS contributed to data acquisition, data interpretation and manuscript writing. CAS critically revised manuscript. JD analysed data and contributed to study intellectual content and manuscript writing. MMSC-S critically revised manuscript.

Funding São Paulo Research Foundation (FAPESP - grants 2008/58238-4, 2012/22997-4, 2014/14880-5 and 2014/50489-9) and Fundação para a Ciência e a Tecnologia, Portugal (Postdoctoral fellowship SFRH / BPD / 111454 / 2015).

Competing interests None declared.

Patient consent for publication Parental/guardian consent obtained.

Ethics approval CAPPesq ethical committee Nº 1085/09 and 195.419.

Provenance and peer review Not commissioned; externally peer reviewed.

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-216410).

To cite Liphaus BL, Caramalho I, Rangel-Santos A, *et al*. *Ann Rheum Dis* 2020;**79**:427–428.

Received 1 October 2019
Revised 9 December 2019
Accepted 11 December 2019
Published Online First 18 December 2019

Ann Rheum Dis 2020;**79**:427–428. doi:10.1136/annrheumdis-2019-216410

ORCID iD
Bernadete L Liphaus http://orcid.org/0000-0002-2624-3233

REFERENCES

1 Carneiro-Sampaio M, Liphaus BL, Jesus AA, *et al*. Understanding systemic lupus erythematosus physiopathology in the light of primary immunodeficiencies. *J Clin Immunol* 2008;28:34–41.
2 Alperin JM, Ortiz-Fernández L, Sawalha AH. Monogenic lupus: a developing paradigm of disease. *Front Immunol* 2018;9:2496.
3 Alkhairy OK, Abolhassani H, Rezaei N, *et al*. Spectrum of phenotypes associated with mutations in LRBA. *J Clin Immunol* 2016;36:33–45.
4 Liphaus BL, Carneiro-Sampaio M. Pediatric immunological disorders Biorepository (PIDBioRep) at University of São Paulo medical school, Brazil. *Biopreserv Biobank* 2013;11:324–6.
5 Sharapova SO, Haapaniemi E, Sakovich IS, *et al*. Novel LRBA mutation and possible germinal mosaicism in a Slavic family. *J Clin Immunol* 2018;38:471–4.
6 Jury EC, Flores-Borja F, Kalsi HS, *et al*. Abnormal CTLA-4 function in T cells from patients with systemic lupus erythematosus. *Eur J Immunol* 2010;40:569–78.

Rapamycin prevents the impairments of social recognition induced by anti-P antibody in a murine model

The antiribosomal P antibody (anti-P) is detected predominantly in patients with systemic lupus erythematosus (SLE)[1] and associated with a variety of neuropsychiatric manifestations (psychosis, mood disorders, cognitive decline, seizures and aseptic meningitis).[2 3] A causal link between anti-P and the neuropsychiatric problems remains to be verified. Recently, anti-P has increasingly been associated with memory impairments. By passive transfer of anti-P into the brain of animals, previous experiments have revealed that hippocampus neurons are prime targets of anti-P, and spatial memory is impaired by the transferred anti-P.[4 5] However, it remains unknown whether anti-P affects on the social memory (the memory of familiar conspecifics). Recently, the ventral CA1 region of hippocampus (vCA1) has been found to play a necessary and sufficient role in social memory.[6] We, therefore, directly injected anti-P IgG (1.7 mg/mL, 0.5 µL) isolated from SLE patient sera, or control IgG from normal individuals or vehicle (artificial cerebrospinal fluid) into vCA1 of normal mice (see details in online supplementary text), and at 24 hours later, we used the social discrimination task to evaluate the impact of anti-P in social memory of mice. As shown in figure 1A, a test mouse was placed in a plexiglass arena, and two pencil-wire cups were placed on opposing corners (one was empty, the other enclosed a mouse). The test mouse habituated to the stimulus mouse during the first three sessions (5 min), rendering it 'familiar'. During the fourth session, a novel mouse was placed in the opposing cup and the three mice were in the same arena. The subject was tested for discrimination between the novel and familiar mouse. Mice received vehicle or control IgG injection showed a longer duration for interaction to a novel mouse than to a familiar mouse, whereas anti-P-injected mice had no preference to a novel mouse, indicating an impairment of social memory (figure 1A,B). We also found that the olfactory and locomotor abilities were not altered in the mice (see online supplementary text and figures S1 and S2), suggesting that the anti-P injection did not cause sensory and motor deficits.

Next, we used a new cohort of mice to examine whether application of rapamycin can protect against the anti-P-induced impairment. The mice were randomly divided into three groups: received daily intraperitoneal injection of rapamycin (0.1 mg/kg) for 7 days before anti-P injection (Rapa +anti-P), the same dose of saline injection and anti-P as the

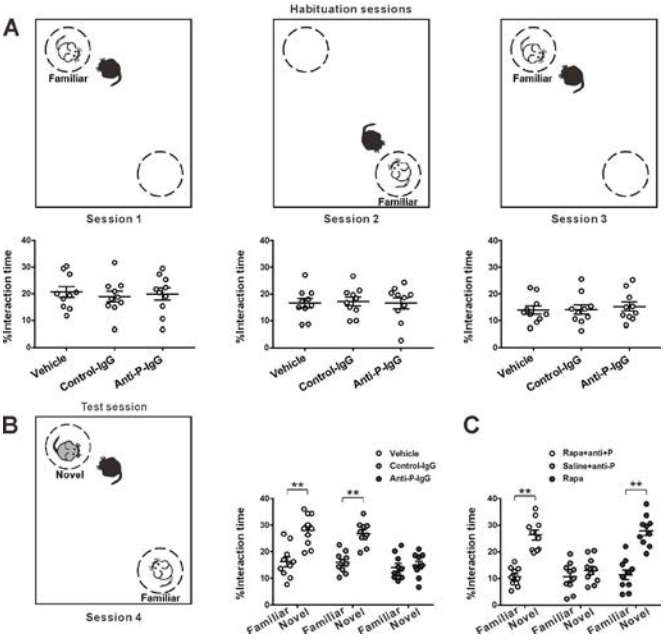


Figure 1 Behavioural schematic, and interaction time when the test mouse was interacting with the stimulus mice in habituation (A) and test sessions (B). Effect of pretreatment of rapamycin. (C) All data are displayed as mean±SEM; n=10. **P<0.01, t-test. anti-P, antiribosomal P.

control group (Saline +anti-P), and rapamycin alone (Rapa). The mice of Rapa +anti-P showed a significant discrimination between the familiar and novel mice, while those of saline +anti-P group did not (figure 1C), suggesting that rapamycin can prevent the social memory impairment induced by anti-P. The application of rapamycin alone has no significant effect on social memory. Here, we present the first evidence showing a detrimental role of anti-P in social memory and the preventive effect of rapamycin. Further systemic experiments are warranted to examine whether and how rapamycin can rescue the autoantibody-induced impairments in patients.

Xuejiao Wang,[1] Pingting Yang,[2] Ling Qin [1]

[1]Department of Physiology, China Medical University, Shenyang, China
[2]Department of Rheumatology and Immunology, First Affiliated Hospital, China Medical University, Shenyang, China

Correspondence to Professor Ling Qin, Department of Physiology, China Medical University, Shen-yang 110122, China; qinlingling@yahoo.comDr Pingting Yang; yangpingtingting@163.com

Handling editor Josef S Smolen

Contributors XW was primarily responsible for experiment procedures, data collection and analysis; PY collected the clinical samples and assisted with manuscript preparation; LQ was primarily responsible for experimental design, data analysis and manuscript preparation. All authors contributed to and approved of the final version of the manuscript.

Funding This work was supported by the grants from National Nature Science Foundation of China under grant (31 671 080 to LQ), Department of Science and Technology of Liaoning Province (2017225024 and 2018225111 to PY) and Department of Science and Technology of Shenyang (19-109-4-15 to PY).

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval All procedures were approved by the Animal Care and Use Committee of China Medical University (No. KT2018060) and strictly adhere to the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health.

Provenance and peer review Not commissioned; externally peer reviewed.

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-216563).

To cite Wang X, Yang P, Qin L. *Ann Rheum Dis* 2020;**79**:428–429.

Received 30 October 2019
Revised 22 November 2019
Accepted 1 December 2019
Published Online First 9 December 2019

Ann Rheum Dis 2020;**79**:428–429. doi:10.1136/annrheumdis-2019-216563

ORCID iD
Ling Qin http://orcid.org/0000-0001-8717-3833

REFERENCES

1 Sato T, Uchiumi T, Ozawa T, *et al*. Autoantibodies against ribosomal proteins found with high frequency in patients with systemic lupus erythematosus with active disease. *J Rheumatol* 1991;18:1681–4.
2 Bonfa E, Golombek SJ, Kaufman LD, *et al*. Association between lupus psychosis and anti-ribosomal P protein antibodies. *N Engl J Med* 1987;317:265–71.
3 Yoshio T, Masuyama J, Ikeda M, *et al*. Quantification of antiribosomal P0 protein antibodies by ELISA with recombinant P0 fusion protein and their association with central nervous system disease in systemic lupus erythematosus. *J Rheumatol* 1995;22:1681–7.
4 Gaburo N, Timo-Iaria C, Timo-Iaria C, *et al*. Electrophysiological dysfunction induced by anti-ribosomal P protein antibodies injection into the lateral ventricle of the rat brain. *Lupus* 2017;26:463–9.
5 Bravo-Zehnder M, Toledo EM, Segovia-Miranda F, *et al*. Anti-Ribosomal P protein autoantibodies from patients with neuropsychiatric lupus impair memory in mice. *Arthritis Rheumatol* 2015;67:204–14.
6 Okuyama T, Kitamura T, Roy DS, *et al*. Ventral CA1 neurons store social memory. *Science* 2016;353:1536–41.

Type I interferon signature predicts response to JAK inhibition in haploinsufficiency of A20

The anti-inflammatory protein A20, encoded by *TNFAIP3*, is a ubiquitin-modifying enzyme that targets proinflammatory molecules, including those upstream of the transcription factor nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB). Patients with heterozygous loss-of-function *TNFAIP3* mutations develop haploinsufficiency of A20 (HA20), a systemic autoinflammatory disease that can cause severe end-organ pathology.[1–3] No available medication directly targets NF-κB signalling; thus, treatment decisions are based on clinical experience. Mild cases are treated with disease-modifying antirheumatic drugs, whereas severe cases are treated with systemic corticosteroids and biological agents, including tumour necrosis factor (TNF)-α and IL-1 receptor (IL-1R) blockade.[1 2] We report that a type I interferon (IFN) signature, or elevation of IFN-stimulated genes (ISGs),

correlates with disease activity and predicts response to janus kinase (JAK) inhibition in HA20.

A cohort of 12 patients with HA20 is followed at the NIH Clinical Center. All patients were diagnosed by Sanger sequencing, and increased NF-κB activity was confirmed with luciferase assay.[2] Five patients had disease that was treatment-refractory or caused end-organ pathology. P1 was a 15-year-old female with p.T604Rfs*93 and severe gastrointestinal ulcerations refractory to TNF-α and IL-1R inhibition. P2–P4 were members of the same extended family with p.F224Sfs*4, aged 28, 32 and 61 years. All three had incomplete responses to TNF-α and IL-1R blockade; P2 had retinal vasculopathy and neuroinflammation, whereas P3 had membranous nephropathy. P5 was an 8-year-old girl with an unreported p.L626Vfs*45 and autoinflammatory liver disease. Notably, T604Rfs*93 and p.L626Vfs*45 are in the fourth zinc finger domain (ZnF4), which mediates several critical functions of A20.[4] Autoinflammatory disease activity index scores ranged from 51 to 117 for P1–P4; P5 was asymptomatic but had severe hepatic inflammation with fibrosis (figure 1A–B).

NF-κB and other A20-regulated signalling molecules can induce type I IFNs,[2] leading us to hypothesise that treatment-refractory HA20 might be characterised by increased ISG expression. Accordingly, we measured expression in the whole blood of all five patients and found elevated ISGs compared with healthy volunteers (figure 1C). As comparators, we also investigated three patients with quiescent HA20; ISGs were not elevated in these subjects (figure 1C). To determine whether NF-κB activation indirectly induced ISGs in HA20 patients, we stimulated healthy and HA20 peripheral blood mononuclear cells in vitro with the NF-κB activating cytokines TNF-α and IL-1β. We were unable to detect *IFNA* or IFN-α in stimulated or unstimulated cells (data not shown). ISG expression decreased over time in unstimulated HA20 cells; for some ISGs, stimulation prevented this reduction. However, neither TNF-α nor IL-1β significantly induced ISGs in healthy or HA20 cells (online supplementary figure S1A–B). This suggests that A20-mediated regulation of NF-κB may induce some ISGs, but that other mechanisms also promote ISG expression in HA20.

Patients with mutations that enhance IFN signalling have been successfully treated with JAK inhibitors, which target the signalling molecules downstream of type I IFNs.[5] This led us to hypothesise that JAK inhibition would be an effective therapeutic strategy for treatment-refractory HA20. Under the IRB-approved protocol 94-HG-0105, we initiated treatment with tofacitinib monotherapy 2.5 mg two times per day for P5, and 5 mg two times per day for P1, P3 and P4. P2 declined tofacitinib. At the time of analysis, treatment duration ranged from 5 to 24 months. Clinical and immunological responses were seen in all four patients (figure 1D–F). Proteinuria in P3 improved from 3+ to 1+ on dipstick analysis, and hepatic transaminases in P5 decreased progressively (figure 1E). Tofacitinib was well-tolerated in all four patients. No opportunistic or severe infections were reported during treatment. Mean haematological parameters and lipid levels remained stable, and there were no cardiovascular or thrombotic events (online supplementary figure S1C–D).

This is the first report that a type I IFN signature correlates with active disease and predicts clinical response in HA20, an autoinflammatory disease that is not a primary interferonopathy. We also expand the spectrum of HA20-associated phenotypes to include severe hepatic inflammation in the absence of systemic features. Together with a recent report of HA20 treatment with the JAK1/2 inhibitor baricitinib, we

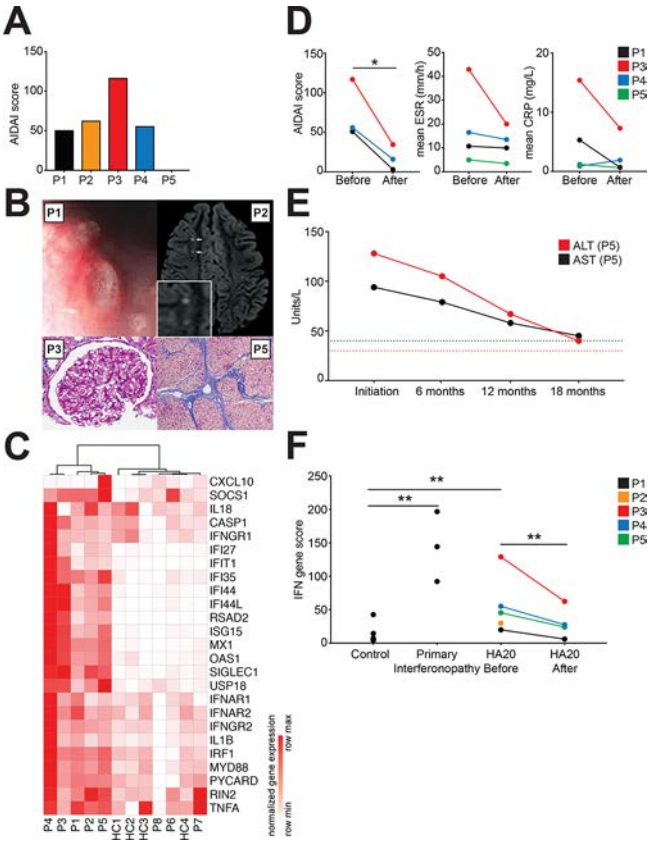


Figure 1 (A) Autoinflammatory disease activity index (AIDAI) scores for five HA20 patients with severe active disease. P1–P4 had AIDAI scores between 51 and 117. P5 was asymptomatic, with an AIDAI score of 0. (B) Organ pathology in five HA20 patients with active disease. Disease manifestations were heterogeneous and included orogenital ulcerations (P1–P4), neuroinflammatory disease (P2), membranous nephropathy (P3, H&E) and autoinflammatory hepatitis with fibrosis (P5, Masson trichrome). (C) Expression of IFN-stimulated genes (ISGs) in HA20 patients and healthy volunteers. Expression is shown for five HA20 patients with active disease (P1–P5), three HA20 patients with quiescent disease (P6–P8) and four healthy volunteers (HC1–HC4). The dendogram shows unbiased hierarchical clustering of the 12 subjects. (D,E) Clinical response of HA20 patients to tofacitinib treatment. (D) Values assessed before and after treatment initiation are shown for AIDAI scores (P1–P3), mean erythrocyte sedimentation rate (ESR) and mean C reactive protein (CRP) (P1–P3, P5). (E) Mean aspartate aminotransferase (AST) and alanine aminotransferase (ALT) levels are shown before and throughout treatment with tofacitinib (P5). (F) Immunological response of HA20 patients to tofacitinib treatment. Geometric mean expression of interferon-stimulated genes are shown in healthy volunteers, patients with primary interferonopathies, HA20 before treatment (P1–P5), and HA20 after treatment (P1–P5). *P<0.05, **p<0.01. Mann-Whitney analysis (for unpaired analysis of healthy volunteers vs HA20), paired t-test (for paired analysis before and after treatment with tofacitinib). HA20, haploinsufficiency of A20.

provide compelling evidence that JAK inhibition is safe and effective for HA20.[6] A20 directly targets NF-κB signalling and the NLRP3 inflammasome; although A20 is also described to regulate IFN signalling, the underlying mechanisms are incompletely characterised.[4 7] Our data suggest that disease activity correlates with elevations of multiple proinflammatory cytokines and increased ISG expression.[2] Future studies will be needed to identify the targets of A20 in various inflammatory

cells, and the mechanisms through which A20 constrains type I IFN responses.

Daniella Muallem Schwartz ,[1] **Sarah A Blackstone,**[1] **Natalia Sampaio-Moura,**[2] **Sofia Rosenzweig,**[2] **Aarohan M Burma,**[1] **Deborah Stone,**[2] **Patrycja Hoffmann,**[2] **Anne Jones,**[2] **Tina Romeo,**[2] **Karyl S Barron,**[1] **Meryl A Waldman,**[3] **Ivona Aksentijevich,**[2] **Daniel L Kastner,**[2] **Joshua D Milner,**[1] **Amanda K Ombrello**[2]

[1]NIAID, National Institutes of Health, Bethesda, Maryland, USA
[2]NHGRI, National Institutes of Health, Bethesda, Maryland, USA
[3]NIDDK, National Institutes of Health, Bethesda, Maryland, USA

Correspondence to Dr Daniella Muallem Schwartz, NIAID, National Institutes of Health, Bethesda, Maryland, USA; Daniella.Schwartz@nih.gov

Handling editor Josef S Smolen

Acknowledgements We thank the patients and their families for participating in the study. We thank G Ben-Yakov for assistance in managing the patients and D Kleiner for high-resolution histology images.

Contributors Conceptualisation: DMS and AKO. Methodology: DMS and AKO. Investigation: DMS, SB, NS-M, AB, SR, DS, PH, TR, AJ, IA and KSB. Formal analysis: DMS and SB. Writing: DMS, SB and AKO. Supervision: DLK, JDM and AKO.

Funding This study was supported by the Intramural Research Programs of the National Human Genome Research Institute, the National Institute for Allergy and Infectious Diseases, the National Institute for Arthritis and Musculoskeletal and Skin Diseases, and the NIH Transitional Program for Clinical Research.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval NIH Institutional Review Board, 94-HG-0105, IRB00000006.

Provenance and peer review Not commissioned; externally peer reviewed.

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-215918).

To cite Schwartz DM, Blackstone SA, Sampaio-Moura N, *et al*. *Ann Rheum Dis* 2020;**79**:429–431.

Received 19 June 2019
Revised 25 October 2019
Accepted 7 November 2019
Published Online First 25 November 2019

Ann Rheum Dis 2020;**79**:429–431. doi:10.1136/annrheumdis-2019-215918

ORCID iD
Daniella Muallem Schwartz http://orcid.org/0000-0002-1660-8438

REFERENCES

1 Aeschlimann FA, Batu ED, Canna SW, *et al*. A20 haploinsufficiency (HA20): clinical phenotypes and disease course of patients with a newly recognised NF-kB-mediated autoinflammatory disease. *Ann Rheum Dis* 2018;77:728–35.
2 Zhou Q, Wang H, Schwartz DM, *et al*. Loss-of-function mutations in TNFAIP3 leading to A20 haploinsufficiency cause an early-onset autoinflammatory disease. *Nat Genet* 2016;48:67–73.
3 Kadowaki T, Ohnishi H, Kawamoto N, *et al*. Haploinsufficiency of A20 causes autoinflammatory and autoimmune disorders. *J Allergy Clin Immunol* 2018;141:1485–8.
4 Das T, Chen Z, Hendriks RW, *et al*. A20/Tumor necrosis factor α-induced protein 3 in immune cells controls development of autoinflammation and autoimmunity: lessons from mouse models. *Front Immunol* 2018;9:104.
5 Sanchez GAM, Reinhardt A, Ramsey S, *et al*. JAK1/2 inhibition with baricitinib in the treatment of autoinflammatory interferonopathies. *J Clin Invest* 2018;128:3041–52.
6 Mulhern CM, Hong Y, Omoyinmi E, *et al*. Janus kinase 1/2 inhibition for the treatment of autoinflammation associated with heterozygous TNFAIP3 mutation. *J Allergy Clin Immunol* 2019;144:863–6.
7 De Wilde K, Martens A, Lambrecht S, *et al*. A20 inhibition of STAT1 expression in myeloid cells: a novel endogenous regulatory mechanism preventing development of enthesitis. *Ann Rheum Dis* 2017;76:585–92.

Risk of malignancy in patients treated for systemic necrotising vasculitis

Reduction in cyclophosphamide cumulative dose and introduction of newer immunosuppressive drugs may reduce the malignant burden of systemic necrotising vasculitis (SNV).[1 2] This study aimed to describe malignancies recorded in five randomised controlled trials in SNV conducted by the French Vasculitis Study Group and to identify predictive factors.

CHUSPAN, CHUSPAN 2, WEGENT, CORTAGE and MAINRITSAN trials evaluated different therapeutic strategies, summarised in online supplementary table S1, for the treatment of newly diagnosed or relapsing SNV. Informations regarding methods and references are provided in the online supplementary materials. The primary endpoint was the occurrence of malignancy.

A total of 733 patients included between 1993 and 2012 were pooled. Baseline characteristics of the population are summarised in table 1. During a 4485.9 person-years (PY) observation period, 39 (5.3%) patients developed malignancies (869.5 per 100 000 PY), including solid cancers in 34 (4.6%) cases (757.9 per 100000 PY) and haematological malignancies in 5 (0.7%) cases (111.5 per 100000 PY). The median interval from inclusion to malignancy's diagnosis was 4.1 (IQR 1.4–8.1) years. The calculated standardised incidence ratios (SIR) for all cancers showed no difference between this cohort and the general population described in the French National registry[3] (SIR 0.95 (0.68–1.30); p=0.84). Solid cancers included gastrointestinal cancers in nine (26.5%) patients, urogenital cancers in eight (23.5%) patients, non-melanoma skin cancers (NMSC) in seven (20.6%) patients, lung cancers in six (17.6%) patients, breast cancers in two (5.9%) patients and brain tumour and carcinoma of unknown primary origin in one patient each. Haematological malignancies included myelodysplastic syndromes in two patients, multiple myeloma, myeloproliferative syndrome and kidney lymphoma in one patient each.

In univariate analysis (table 1), malignancy's occurrence was associated with age ≥65 years (HR 2.89 (1.41 to 5.92); p=0.004), estimated glomerular filtration rate <30 mL/min/1.73 m^2 (HR 2.55 (1.17 to 5.55); p=0.019), baseline positive antineutrophil cytoplasmic antibodies (HR 2.79 (1.14 to 6.81); p=0.024), the occurrence of ≥1 relapse(s) during follow-up (HR 2.05 (1.02 to 4.12); p=0.043), the use of azathioprine (HR 5.31 (2.03 to 13.91); p<0.001) and methotrexate (HR 5.11 (1.88–13.89); p=0.001) as maintenance therapy (always used after cyclophosphamide-based induction therapy) and the maintenance therapy duration (HR 1.29 (1.06 to 1.58) p=0.013). In contrast, the combination of cyclophosphamide as induction and rituximab as maintenance therapy was not significanltly associated with malignancy (HR 2.65 (0.48 to 14.64); p=0.262). In multivariate analysis, age ≥65 years (HR 2.38 (1.13 to 4.99); p=0.022), the use of azathioprine (HR 3.05 (1.10 to .42); p=0.032) and methotrexate (HR 3.24 (1.07 to 9.81); p=0.038) as maintenance remained independently associated with the occurrence of malignancy. Similar results were obtained after exclusion of NMSC.

Malignancy was associated with a poorer overall survival (median 12.1 years vs not reached; (p=0.004)). The cause of death was directly related to the malignancy or its treatment in all 19 (100%) patients with cancer who died.

While the association between advanced age and cancer was expected,[4] data regarding the impact of newer therapeutic strategies remain scarce. In our population, the use of conventional

Table 1 Comparison of baseline characteristics according to cancer status

	Overall n=736	No malignancy n=694	Malignancy n=39	HR (95% CI)	P value
Demography					
Age ≥65 years	294 (39.9)	274 (39.5)	20 (51.3)	2.89 (1.41 to 5.92)	**<0.01**
Male gender	399 (54.7)	377 (54.5)	22 (57.9)	1.15 (0.56 to 2.35)	0.71
Inclusion period				1.54 (0.91 to 2.6)	0.11
1993–1999	196 (26.7)	181 (26.1)	15 (38.5)		
2000–2005	238 (32.5)	223 (32.1)	15 (38.5)		
2006–2012	299 (40.8)	290 (41.8)	9 (23.1)		
BVAS	11.5 (11.2)	11.2 (11.1)	15.5 (12.6)	1.01 (0.98 to 1.04)	0.65
FFS ≥1	434 (59.2)	402 (57.9)	32 (82.1)	2.12 (0.91 to 4.93)	0.08
SNV diagnoses					
EGPA	186 (25.4)	184 (26.5)	2 (5.1)	0.24 (0.06 to 1.02)	0.05
GPA	224 (30.6)	209 (30.1)	15 (38.5)	1.72 (0.86 to 3.45)	0.13
MPA	238 (32.5)	220 (31.7)	18 (46.2)	0.49 (0.14 to 1.70)	0.26
PAN	85 (11.6)	81 (11.7)	4 (10.3)	1.55 (0.75 to 3.19)	0.24
Baseline positive ANCA	469 (64.0)	439 (63.3)	30 (76.9)	2.79 (1.15 to 6.81)	**0.02**
PR3	187 (25.5)	176 (25.4)	11 (28.2)	1.33 (0.64 to 2.77)	0.44
MPO	243 (33.2)	226 (32.6)	17 (43.6)	1.83 (0.9 to 3.75)	0.10
Clinical manifestations					
Cutaneous involvement	325 (44.3)	313 (45.1)	12 (30.8)	0.55 (0.26 to 1.16)	0.12
Ocular involvement	96 (13.1)	91 (13.1)	5 (12.8)	1.09 (0.42 to 2.84)	0.86
ENT involvement	370 (50.5)	352 (50.7)	18 (46.2)	1.19 (0.59 to 2.39)	0.63
Pulmonary involvement	445 (60.7)	427 (61.5)	18 (46.2)	0.61 (0.3 to 1.22)	0.16
Cardiovascular involvement	142 (19.4)	134 (19.3)	8 (20.5)	0.57 (0.2 to 1.64)	0.30
Gastrointestinal involvement	147 (20.1)	138 (19.9)	9 (23.1)	1.2 (0.53 to 2.69)	0.66
Neurological involvement	402 (54.8)	378 (54.5)	24 (61.5)	1.23 (0.61 to 2.5)	0.56
Renal involvement	368 (50.2)	341 (49.1)	27 (69.2)	1.66 (0.79 to 3.46)	0.18
eGFR <30 mL/min/1.73 m^2	125 (17.1)	116 (16.7)	9 (23.1)	2.55 (1.71 to 5.55)	**0.02**
Patients with ≥1 relapse(s)	239 (32.6)	219 (31.6)	20 (51.3)	2.05 (1.02 to 4.12)	**0.04**
Patients with ≥1 SI	148 (20.2)	137 (19.7)	11 (28.2)	1.46 (0.06 to 2.99)	0.38
Induction regimen					
GCs alone	161 (22)	157 (22.6)	4 (10.3)	Ref	–
GCs+AZA	78 (10.6)	78 (11.2)	0 (0.0)	–	–
GCs+CYC	494 (67.4)	459 (66.1)	35 (89.7)	2.04 (0.71 to 5.85)	0.18
CYC cumulative dose (g)	8.62 (4.36)	8.50 (4.36)	10.08 (4.07)	0.99 (0.90 to 1.09)	0.84
Maintenance regimen					
No maintenance	410 (55.9)	396 (57.1)	14 (35.9)	Ref	–
MTX	64 (8.7)	55 (7.9)	9 (23.1)	5.11 (1.88 to 13.89)	**<0.01**
AZA	202 (27.6)	188 (27.1)	14 (35.9)	5.31 (2.03 to 13.91)	**<0.01**
RTX	57 (7.8)	55 (7.9)	2 (5.1)	2.65 (0.48 to 14.64)	0.26
Maintenance duration (months)	16.93 (4.1)	16.80 (4.2)	18.96 (1.7)	1.29 (1.06 to 1.58)	**0.01**

Cancer-free survival analysis (Cox model).
Bold indicates statistical significance.
Data are number (%) or mean±SD unless otherwise indicated. Percentages were calculated based on the number of available data.
ANCA, antineutrophil cytoplasmic antibodies; AZA, azathioprine; BVAS, Birmingham Vasculitis Activity Score; eGFR, estimated glomerular filtration rate; EGPA, eosinophilic granulomatosis with polyangiitis; ENT, ear nose and throat; FFS, Five-Factor Score; GCs, glucocorticoids; GPA, granulomatosis with polyangiitis; MPA, microscopic polyangiitis; MPO, myeloperoxidase; MTX, methotrexate; PAN, polyarteritis nodosa; PR3, proteinase 3; RTX, rituximab; SI, severe infection.

immunosuppressive regimens was still associated with a higher risk of malignancy. In contrast, patients treated with rituximab were less likely to develop malignancy, corroborating a recent study, suggesting a safer tolerance profile.[2]

Lastly, a key finding of our study was the lower incidence of malignancy than previously reported in patients with SNV,[1] which was now comparable to that of the general population. This decreased risk of malignancy may be driven by the more extensive use of cyclophosphamide-sparing strategies.[1] Indeed, patients included in this study received a reduced cyclophosphamide cumulative exposure in comparison with those analysed by Heijl *et al* (4.32±1.63 vs 11.8±2.4 months, p<0.0001)[1] or Le Guenno *et al* (8.62±4.36 vs 25.1±38 g, p<0.0001).[5] Taken together, these findings suggest that malignancy should no longer be considered as a predominant feature driving the therapeutic strategy.

Antoine Lafarge ,[1,2] **Adrien Joseph,**[3] **Christian Pagnoux,**[4] **Xavier Puechal,**[1,2] **Pacal Cohen,**[1,2] **Maxime Samson** ,[5] **Mohamed Hamidou,**[6] **Alexandre Karras,**[7] **Matthieu Groh,**[1] **Thomas Quemeneur,**[8] **Camillo Ribi,**[9] **Luc Mouthon,**[1,2,10] **Loic Guillevin,**[1,2,10] **Benjamin Terrier** ,[1,2,10] **on behalf of the French Vasculitis Study Group (FVSG)**

[1]Department of Internal Medicine, Cochin Hospital, Paris, France

[2]National Referral Center for Rare Systemic Autoimmune Diseases, Cochin Hospital, Paris, France
[3]U1138, INSERM, Paris, France
[4]Department of Rheumatology, Mount Sinai Hospital, Toronto, Ontario, Canada
[5]Department of Internal Medicine and Clinical Immunology, Hôpital François Mitterrand, Dijon, France
[6]Department of Internal Medicine, Hôtel Dieu, Nantes, France
[7]Department of Nephrology, Hôpital Européen Georges Pompidou, Paris, France
[8]Department of Internal Medicine, Hôpital de Valenciennes, Valenciennes, France
[9]Department of Immunology, CHUV, Lausanne, Switzerland
[10]Paris Descartes University, Paris, France

Correspondence to Dr Antoine Lafarge, Internal Medicine, Hopital Cochin, Paris 75014, France; antoinelafarge@outlook.comPr Benjamin Terrier; benjamin.terrier@aphp.fr

Correction notice This article has been corrected since it published Online First. The first author's name has been corrected.

Handling editor Josef S Smolen

Contributors LA and BT devised the project, the main conceptual ideas and proof outline, pooled and analysed data from the five randomised controlled trials (RCTs). AJ performed the statistical analysis. CP, XP, PC, CR, LG and BT designed and published the five RCTs involved in this article. LA and BT, wrote the manuscript. All authors discussed the results and contributed to the final manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-216452).

To cite Lafarge A, Joseph A, Pagnoux C, *et al*. *Ann Rheum Dis* 2020;**79**:431–433.

Received 9 October 2019
Revised 12 November 2019
Accepted 15 November 2019
Published Online First 25 November 2019

Ann Rheum Dis 2020;**79**:431–433. doi:10.1136/annrheumdis-2019-216452

ORCID iDs
Antoine Lafarge http://orcid.org/0000-0002-2481-5728
Maxime Samson http://orcid.org/0000-0002-6547-232X
Benjamin Terrier http://orcid.org/0000-0001-6612-7336

REFERENCES

1 Heijl C, Harper L, Flossmann O, *et al*. Incidence of malignancy in patients treated for antineutrophil cytoplasm antibody-associated vasculitis: follow-up data from European vasculitis Study Group clinical trials. *Ann Rheum Dis* 2011;70:1415–21.
2 van Daalen EE, Rizzo R, Kronbichler A, *et al*. Effect of rituximab on malignancy risk in patients with ANCA-associated vasculitis. *Ann Rheum Dis* 2017;76:1064–9.
3 Binder-Foucard F, Bossard N, Delafosse P, *et al*. Cancer incidence and mortality in France over the 1980-2012 period: solid tumors. *Rev Epidemiol Sante Publique* 2014;62:95–108.
4 López-Otín C, Blasco MA, Partridge L, *et al*. The hallmarks of aging. *Cell* 2013;153:1194–217.
5 Le Guenno G, Mahr A, Pagnoux C, *et al*. Incidence and predictors of urotoxic adverse events in cyclophosphamide-treated patients with systemic necrotizing vasculitides. *Arthritis Rheum* 2011;63:1435–45.

Rituximab for chronic periaortitis without evidence of IgG4-related disease: a long-term follow-up study of 20 patients

Chronic periaortitis (CP) is a rare condition characterised by a peri-aortoiliac fibro-inflammatory tissue. A total of 20%–50% of the cases are immunoglobulin G4 (IgG4)-related, based on histological evidence of IgG4+ plasma cell infiltration (on a background of dense lymphoplasmacytic infiltrates, storiform fibrosis and tissue eosinophilia) and/or increased serum IgG4.[1]

Glucocorticoids are the first-line therapy for CP.[2] However, some patients are refractory, frequently relapsing or have contraindications to glucocorticoids. The anti-CD20 monoclonal antibody rituximab proved efficacious in systemic forms of IgG4-related disease (IgG4-RD) including IgG4-related CP,[3] but data on IgG4-unrelated CP are scarce.[4–6] In this study, we tested rituximab in CP patients without evidence of IgG4-RD who had relapsing/refractory disease or contraindications to standard-dose glucocorticoids.

We included patients with active, IgG4-unrelated CP who received rituximab (October 2009 to April 2017). Online supplementary methods describe the diagnostic and follow-up procedures, the definitions of remission and refractory, and the statistical analysis.

Twenty consecutive patients were included. Two of them were previously reported.[5] Of the eight patients with available CP biopsies, none had significant $IgG4^+$-plasma cell infiltration. None had other biopsy-proven IgG4-related lesions or high serum IgG4. Four patients were newly diagnosed and had contraindications to standard-dose glucocorticoids, 13 were frequent relapsers and 3 refractory.

The patients' clinical manifestations are reported in supplementary table S1. Rituximab (1000 mg 2 weeks apart or 375 mg/m^2/week×4 weeks) was given alone to 4 patients and with prednisone (median initial dose 25 mg/day, IQR 25–50 mg/day) to 16; six received rituximab maintenance (single 1000 mg doses every 6–8 months).

At month 6, all patients were symptom-free; erythrocyte sedimentation rate (ESR) (figure 1) and C-reactive protein (CRP) dropped (respectively, $p<0.0001$ and $p=0.01$, vs baseline). The proportion of patients with ureteral involvement decreased from 65% to 40% (supplementary table S2). A significant reduction was observed in periaortic ($p=0.001$) and peri-iliac ($p=0.006$) CP thickness, maximum standardised uptake value ($p=0.0001$) and prednisone dose ($p<0.0001$) (figure 1). Supplementary figure S1 shows representative CT/positron emission tomography (PET) responses, and supplementary figure S2 shows PET uptake grades after treatment. At month 12, two patients were lost to follow-up. The remaining 18 were asymptomatic, 11 (61%) being glucocorticoid-free. Ureteral involvement rate, ESR, CRP and CP thickness further decreased (supplementary table S2, figure 1). At month 18, all the 16 assessable patients were asymptomatic; two (12.5%) had ureteral involvement. Fluorodeoxyglucose (FDG) uptake further declined ($p<0.0001$ vs baseline).

During the follow-up (median 38 months, IQR 17–61 onths), 15 patients (75%) achieved remission; of them, three relapsed (months 4, 47 and 59) (supplementary figures S3 and S4). Two were successfully retreated with rituximab and one received methotrexate +prednisone. The main outcomes did not significantly differ between the rituximab monotherapy and the rituximab +prednisone treatment regimens, and between those who received protocolised rituximab retreatment and those who did not (data not shown), although these subgroups were small for reliable comparisons.

One patient died for stroke (month 12) and another developed chronic lymphocytic leukaemia (month 33). The remaining adverse events were graded 1–3 (supplementary table S3).

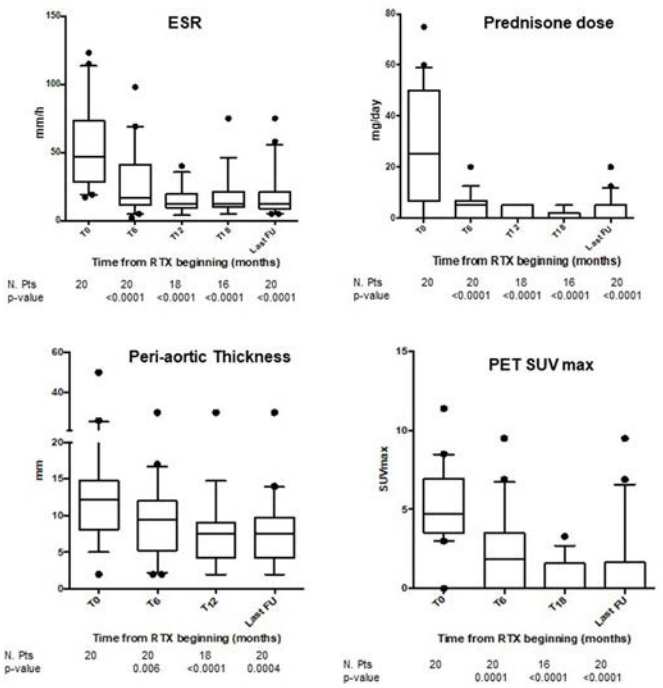


Figure 1 Variation in erythrocyte sedimentation rate (ESR), prednisone dose, periaortic thickness of chronic periaortitis (CP) and SUVmax at different time points. Significant reductions in ESR, prednisone dose, periaortic CP thickness and SUVmax were observed after rituximab (RTX) therapy. These variations were analysed using Wilcoxon signed-rank test. The reported p values refer to the comparisons between each time point and baseline. Significant reductions were also observed between T12 and T6 for ESR (p=0.02), prednisone dose (p=0.007) and CP thickness (p=0.001). Data are shown as box plots. Each box represents 25th–75th percentiles. Lines inside the boxes indicate the median. The whiskers represent 10th–90th percentiles. Circles indicate outliers. T0, T6, T12, T18 and Last FU denote, respectively, the time of RTX therapy, month 6, month 12, month 18 and last follow-up. PET, positron emission tomography; SUVmax, maximum standardised uptake value.

Our results show that rituximab achieves objective and metabolic responses and a high remission rate in CP patients without IgG4-RD who had relapsing/refractory disease or contraindications to standard-dose glucocorticoids. Although responses early after rituximab could also result from concomitant glucocorticoid therapy, further improvement on CT/MRI and PET was also seen after month 12, when most patients had discontinued glucocorticoids. Rituximab was well-tolerated.

Our study was retrospective and had a small sample size, which also limited subgroup comparisons (eg, rituximab alone vs rituximab plus glucocorticoids, and protocolised vs on-demand re-retreatment). Another limitation relates to the lack of biopsies in 60% of the patients, which makes their IgG4-unrelatedness uncertain. However, biopsies are routinely performed in only 10%–30% of CP patients.[2 4] These drawbacks withstanding, our findings encourage rituximab use for difficult-to-treat CP and advocate larger confirmatory studies.

Maria Letizia Urban,[1] Federica Maritati,[1] Alessandra Palmisano,[1] Paride Fenaroli,[2] Francesco Peyronel,[2] Giorgio Trivioli,[3] Stefania Ferretti,[1] Consuelo De Biase,[4] Peter C Grayson ,[5] Francesco Pegoraro,[6] Domenico Prisco,[7] Paola Romagnani,[8] Giacomo Emmi,[1] Augusto Vaglio [9]

[1]Department of Experimental and Clinical Medicine, University of Florence, Firenze, Italy
[2]Nephrology, Parma University Hospital, Parma, Italy, Italy
[3]Nephrology Unit, Department of Clinical and Experimental Medicine, University of Parma, Parma, Italy
[4]Nephrology, University of Turin, Torino, Italy
[5]Pediatrics, National Institutes of Health/NIAMS, Bethesda, Maryland, USA
[6]Azienda Ospedaliero Universitaria Ospedale Pediatrico Meyer, Firenze, Italy
[7]Department of Experimental and Clinical Medicine, University of Florence, AOU Careggi, Florence, Italy
[8]Nephrology and Dialysis Unit, Meyer Children's University Hospital, Florence, Italy
[9]Biomedical Experimental and Clinical Sciences, University of Florence, Firenze, Italy

Correspondence to Dr Augusto Vaglio, University of Florence, 50121 Firenze, Italy; augusto.vaglio@virgilio.it

Handling editor Josef S Smolen

Contributors MLU, FM and AV designed the study, analysed the data and drafted the manuscript. MLU, FM, AP, PF, F Peyronel, GT, SF, CDB, PCG, GE and AV followed the patients and collected the data. F Pegoraro and MLU generated the figures. PCG, DP, PR and GE critically reviewed the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement statement Given the retrospective nature of this work, it was not possible to involve patients or the public in the design, or conduct, or reporting, or dissemination of this research.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

► Additional material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2019-216258).

To cite Urban ML, Maritati F, Palmisano A, *et al*. *Ann Rheum Dis* 2020;**79**:433–434.

Received 3 September 2019
Revised 24 October 2019
Accepted 5 November 2019
Published Online First 19 November 2019

Ann Rheum Dis 2020;**79**:433–434. doi:10.1136/annrheumdis-2019-216258

ORCID iDs
Peter C Grayson http://orcid.org/0000-0002-8269-9438
Augusto Vaglio http://orcid.org/0000-0002-3814-9172

REFERENCES

1 Rossi GM, Rocco R, Accorsi Buttini E, *et al*. Idiopathic retroperitoneal fibrosis and its overlap with IgG4-related disease. *Intern Emerg Med* 2017;12:287–99.
2 Vaglio A, Palmisano A, Alberici F, *et al*. Prednisone versus tamoxifen in patients with idiopathic retroperitoneal fibrosis: an open-label randomised controlled trial. *The Lancet* 2011;378:338–46.
3 Ebbo M, Grados A, Samson M, *et al*. Long-Term efficacy and safety of rituximab in IgG4-related disease: data from a French nationwide study of thirty-three patients. *PLoS One* 2017;12:e0183844.
4 Wallwork R, Wallace Z, Perugino C, *et al*. Rituximab for idiopathic and IgG4-related retroperitoneal fibrosis. *Medicine* 2018;97:e12631.
5 Maritati F, Corradi D, Versari A, *et al*. Rituximab therapy for chronic periaortitis. *Ann Rheum Dis* 2012;71:1262–4.
6 Maritati F, Rocco R, Accorsi Buttini E, *et al*. Clinical and prognostic significance of serum IgG4 in chronic periaortitis. An analysis of 113 patients. *Front Immunol* 2019;10.

Catching the falling star: points to consider when using propensity scores

Bergstra *et al* conducted a sophisticated study of a pertinent clinical question, 'How to treat patients with rheumatoid arthritis when methotrexate has failed?'[1] In particular, I welcome the introduction of the 'multiple' propensity score (PS) (better known by its more established name: generalised propensity score[2 3]) to observational studies in rheumatology. The authors provided a clear step-by-step tutorial which I am sure will be frequently used and cited. Since their article also aimed to teach, I would like to highlight some important points to readers who consider replicating these methods.

Use of PS methods has grown exponentially in clinical research. Suboptimal practice has unfortunately led to a level of scepticism for these versatile techniques.[4] The theory and assumptions are often more involved, but some points can be easily conceptualised using its underlying causal aim: emulating a randomised controlled trial (RCT).

In an RCT, both measured and unmeasured confounders are equally distributed between the treatment arms. What is not known at randomisation is each patient's future outcome (but is known to the analyst of observational data). Variable selection for the PS model should be based on clinical knowledge and prior literature, rather than regressing on the outcome. There are many downsides to base variable selection on statistical significance, such as the dependence on sample size (which was relatively small in this study, n=509). There are two popular schools of thought for variable selection: variables related to both exposure and outcome, or strongly related to the outcome, should be selected; variables that are only strongly associated with the exposure should not.[5] Alternatively, all available baseline variables can be included.[6]

Second, as the authors alluded to, no patients in an RCT would have zero probability of receiving any treatment; extrapolating results to patients who would never have received the treatment is a danger when using observational data. The authors removed such patients. However, their PS distribution still showed a very high proportion of patients in the biologic DMARD ±conventional synthetic DMARD(s) group with very small probably (<0.05) of receiving the treatment. A better-established method is to trim the PS, for example, by centiles.[7] Doing so can limit generalisability but ensures validity of comparison. As a side note, it should be clarified that this analysis assumed no indication bias within each of the three treatment groups (ie, clinicians randomly chose each combination of drugs within each group).

Third, I would like to highlight to the readership that generalised PS, like binary PS, can also be used by matching or weighting.[2] Unlike adjusting for PS in the outcome model, these methods have clear estimands (average treatment effect or average treatment effect of the treated) and allow validated diagnostics of balance, such as using standardised mean difference. Adjusting for PS is generally not recommended (except as a method of variable reduction) since it requires additional assumptions of the relationship between the PS and outcome.[6]

The authors offered a reason against alternative PS methods in that they may result in a smaller sample size. This is true. But such restrictions are essential to ensure comparability. If indeed a large proportion of cases were lost when they tried to match, then the published sample would undoubtedly remain afflicted by indication bias. It was surprising that the adjusted and unadjusted results were the same in the face of clearly different indications for each treatment arm.

Finally, results from 1.3% of patients of the original study population (509 selected from 37 808) should be interpreted with caution. Even more so since the authors diligently reported that the included and excluded patients were clearly different. As a testament to the versatility of PS-based methods, a sensitivity analysis could be performed using 'the propensity for missingness'—better known as inverse probability of censoring weights.[8] In response to Dr Ahmed[4] : the PS is far from a falling star when appropriately used.

Fengchen Ouyang

Retired, Xi'an, Shaanxi, China

Correspondence to Fengchen Ouyang, Retired, Xi'an, Shannxi, China; oyfc18@gmail.com

Contributors None

Competing interests None declared.

Patient consent Not required.

Provenance and peer review Not commissioned; internally peer reviewed.

To cite Ouyang F. *Ann Rheum Dis* 2019;**79**:e26.

Received 4 December 2018
Accepted 8 December 2018
Published Online First 30 December 2018

Linked

► https://doi.org/10.1136/annrheumdis-2018-214892

Ann Rheum Dis 2019;**79**:e26. doi:10.1136/annrheumdis-2018-214872

REFERENCES

1 Bergstra SA, Winchow LL, Murphy E, *et al*. How to treat patients with rheumatoid arthritis when methotrexate has failed? The use of a multiple propensity score to adjust for confounding by indication in observational studies. *Ann Rheum Dis* 2018. doi: 10.1136/annrheumdis-2018-213731 [Epub ahed of print: 16 Oct 2018].
2 Austin PC. Assessing the performance of the generalized propensity score for estimating the effect of quantitative or continuous exposures on binary outcomes. *Stat Med* 2018;37:1874–94.
3 Imbens G. The role of the propensity score in estimating dose-response functions. *Biometrika* 2000;87:706–10.
4 Ahmed S. Metering the METEOR in methotrexate failure: is propensity score a falling star? *Ann Rheum Dis* 2018. doi: 10.1136/annrheumdis-2018-214612 [Epub ahed of print: 8 Nov 2018].
5 Brookhart MA, Schneeweiss S, Rothman KJ, *et al*. Variable selection for propensity score models. *Am J Epidemiol* 2006;163:1149–56.
6 Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 2011;46:399–424.
7 Brookhart MA, Wyss R, Layton JB, *et al*. Propensity score methods for confounding control in nonexperimental research. *Circ Cardiovasc Qual Outcomes* 2013;6:604–11.
8 Hernan M, Robins J. *Causal inference [Internet]*. Boca Raton: Chapman & Hall/CRC. https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/

Response to: 'Catching the falling star: points to consider when using propensity scores' by Ouyang *et al*

We would like to thank Dr Ouyang[1] for his thoughtful response to our article, including additional suggestions regarding the use of the multiple (or generalised) propensity score (PS), which we are sure will be helpful for other researchers who plan to use this technique.[2] We agree with the mentioned suggestions and will provide some additional clarifications.

Like virtually every method used to adjust for bias in observational data, a multiple PS is dependent on the ignorability assumption and assumes that all important confounders are measured and included. Although in practice it is probably impossible to know all confounders, this emphasises the importance of variable selection for the PS model.

We agree with Dr Ouyang that variables should not be included based on statistical significance only (thus not taking into account prior knowledge). Therefore, we followed the approach of Brookhart *et al*[3] and selected all potential variables associated with the outcome of the study based on clinical knowledge. We then based further inclusion on statistical significance, using a conservative p value of <0.10. Although inclusion based on statistical significance has disadvantages, it can be helpful in a situation in which a large number of pretreatment variables is available, especially if the association with the outcome is unclear.[4]

Next, we fully agree that trimming patients without a chance of receiving each treatment of interest (untied observations) is important, and we want to thank Dr Ouyang with the introduction of an alternative approach: to trim the PS on centiles.

Furthermore, previous literature has indeed shown that using a PS for matching is more succesful in reducing bias than using a PS for stratification or covariate adjustment.[5 6] Whereas in our study with three treatment groups, we might have decided to apply matching, in studies with an increasing number of exposure groups matching patients may truly become infeasible. We believe that in such a situation covariate adjustment can be an alternative solution.

Lastly, Dr Ouyang emphasised the risk of selection bias in our study. We used data from METEOR, a daily practice database with as only inclusion criterion that patients have a clinical diagnosis of rheumatoid arthritis (RA). At the time of analysis, this database indeed included 37 808 patients, but for our analysis, we first made a selection of eligible patients (as described, newly diagnosed patients, starting with methotrexate monotherapy and so on), which resulted in a selection of 1561 patients. Next, we compared baseline characteristics of the 509 included and the 1059 eligible but non-included patients. Nevertheless, we acknowledge that a risk of selection bias is still present in our study and should be kept in mind when interpreting results.

Overall, we hope that our article, with the additional suggestions made by the Dr Ouyang, will contribute to the introduction and appropriate use of the multiple (or generalised) PS to observational research in rheumatology.

Sytske Anne Bergstra,[1] Cornelia F Allaart,[1] Robert B M Landewé[2,3]

[1]Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
[2]Amsterdam Rheumatology & Immunology Center, Amsterdam, The Netherlands
[3]Zuyderland Medical Center, Heerlen, The Netherlands

Correspondence to Sytske Anne Bergstra, Department of Rheumatology, Leiden University Medical Center, Leiden 2333 ZA, The Netherlands; s.a.bergstra@lumc.nl

Handling editor Josef S Smolen

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Commissioned; internally peer reviewed.

To cite Bergstra SA, Allaart CF, Landewé RBM. *Ann Rheum Dis* 2020;**79**:e27.

Received 14 December 2018
Revised 17 December 2018
Accepted 17 December 2018
Published Online First 30 December 2018

Linked

► https://doi.org/10.1136/annrheumdis-2018-214872

Ann Rheum Dis 2020;**79**:e27. doi:10.1136/annrheumdis-2018-214892

REFERENCES

1 Ouyang F. Catching a falling star: points to consider when using propensity scores. *Ann Rheum Dis* 2018.
2 Bergstra SA, Winchow LL, Murphy E, *et al*. How to treat patients with rheumatoid arthritis when methotrexate has failed? The use of a multiple propensity score to adjust for confounding by indication in observational studies. *Ann Rheum Dis* 2018. doi: 10.1136/annrheumdis-2018-213731. [Epub ahead of print 16 Oct 2018].
3 Brookhart MA, Schneeweiss S, Rothman KJ, *et al*. Variable selection for propensity score models. *Am J Epidemiol* 2006;163:1149–56.
4 Spreeuwenberg MD, Bartak A, Croon MA, *et al*. The multiple propensity score as control for bias in the comparison of more than two treatment arms: an introduction from a case study in mental health. *Med Care* 2010;48:166–74.
5 Austin PC, Grootendorst P, Anderson GM. A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: a Monte Carlo study. *Stat Med* 2007;26:734–53.
6 Austin PC, Mamdani MM. A comparison of propensity score methods: a case-study estimating the effectiveness of post-AMI statin use. *Stat Med* 2006;25:2084–106.

Role of linoleic acid in autoimmune disorders: a Mendelian randomisation study

I read with great interest the article by Zhao and Schooling[1] regarding the role of linoleic acid in autoimmune disorders. This Mendelian randomisation (MR) analysis suggests that linoleic acid protects against rheumatoid arthritis (RA). However, it has a methodological issue. The choice of the genetic instrumental variables (IV) is essential for a successful MR study. MR analyses using multiple genetic variants can be viewed as a meta-analysis of the causal estimates from each variant.[2] The availability of estimates of both the gene-risk factor and the gene-outcome associations for each of these variants is important. However, the authors used limited numbers of IVs (three single nucleotide polymorphisms (SNP) with top significance and seven SNPs on functionally relevant genes).[1] Genetic instruments tend to have weak power due to the limited availability of population-specific information on genetic associations.[3] Bias from weak instruments can result in misleading estimates of causal effects. If the variants in total explain a larger proportion of the variance in the exposure, this will lead to more precise estimates of causal effects, thus increasing the power for MR analysis.[3] Therefore, the approach of using multiple genetic variants in different gene regions is suitable for an MR study. I applied a two-sample MR analysis using the inverse-variance weighted (IVW), MR-Egger regression and weighted median methods to the data from a genome-wide association study (GWAS) of n-6 polyunsaturated fatty acid (PUFA) metabolism in 8631 adults[4] as an exposure variable and RA GWAS (14 361 cases and 43 923 controls)[5] as an outcome. I selected the independent association of 75 SNPs associated with PUFA metabolism based on a linkage disequilibrium R^2 of 0.001, clumping distance of 10 000 kb and a p value threshold of 5.00E-08 (genome-wide significance). The MR estimates determined using the IVW, weighted median and MR-Egger regression analyses were consistent and do not support a causal inverse association between linoleic acid and the occurrence of RA (beta=0.00008, SE=0.001, p=0.949). The MR-Egger regression revealed that directional pleiotropy was unlikely to have biased the results, and the funnel plot test revealed a symmetry, indicating no evidence of pleiotropy. Including more instruments, where each instrument explains an extra variation in the phenotype, should provide more information on the causal estimate. Thus, I believe that the findings of this MR study should be interpreted by taking the aforementioned methodological concerns into consideration.

Young Ho Lee

Department of Rheumatology, Korea University Medical Center, Seoul 02841, Republic of Korea

Correspondence to Professor Young Ho Lee, Department of Rheumatology, Division of Rheumatology, Korea University Medical Center, Seoul 02841, Republic of Korea; lyhcgh@korea.ac.kr

Handling editor Josef Smolen

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent Not required.

Provenance and peer review Not commissioned; internally peer reviewed.

To cite Lee YH. *Ann Rheum Dis* 2020;**79**:e28.

Received 23 November 2018
Accepted 25 November 2018
Published Online First 18 December 2018

Linked

► https://doi.org/10.1136/annrheumdis-2018-214830

Ann Rheum Dis 2020;**79**:e28. doi:10.1136/annrheumdis-2018-214810

REFERENCES

1 Zhao JV, Schooling CM. Role of linoleic acid in autoimmune disorders: a Mendelian randomisation study. *Ann Rheum Dis* 2018. doi: 10.1136/annrheumdis-2018-214519. [Epub ahead of print: 8 Nov 2018].
2 Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* 2013;37:658–65.
3 Burgess S, Thompson SG, CRP CHD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol* 2011;40:755–64.
4 Guan W, Steffen BT, Lemaitre RN. *Genome-wide association study of plasma n6 polyunsaturated fatty acids within the CHARGE consortium*. Circulation: Genomic and Precision Medicine, 2014.
5 Okada Y, Wu D, Trynka G, *et al*. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* 2014;506:376–81.

Response to: 'Role of linoleic acid in autoimmune disorders: aMendelian randomisation study' by Lee *et al*

We are pleased that our article on the role of linoleic acid (LA) in autoimmune disorders is of interest to readers. However, regarding the methodological issues raised by Lee,[1] several points need to be considered and clarified.

First, Mendelian randomisation (MR) requires stringent assumptions, that is, the genetic instruments are associated with the exposure, are not linked with the outcomes other than via effects on the exposure and no confounders of the associations of the genetic instruments with the outcome exist.[2] We agree that weak instruments which violate these assumptions would lead to biased associations. As such, we are very cautious in the selection of genetic instruments. Specifically, we used the most significant three uncorrelated ($r^2<0.01$) single-nucleotide polymorphisms (SNPs) from a genome-wide association study (GWAS),[3] as previously,[4] and replicated using uncorrelated SNPs in genes relevant to the metabolism of n−6 PUFA, that is, *FADS1*, *FADS2* and *NTAN1*.[5] To ensure the SNPs predicting LA were not confounded, we assessed their Bonferroni corrected associations with key confounders, that is, socioeconomic position (job and Townsend Index) and lifestyle factors (alcohol and smoking), in the UK Biobank. To ensure the selected SNPs were solely linked with autoimmune disorders via effects on LA (no pleiotropy), we checked using three comprehensive curated genetic cross-reference systems, Ensembl (http://www.ensembl.org/index.html), the GWAS catalogue (https://www.ebi.ac.uk/gwas/) and PhenoScanner (www.phenoscanner.medschl.cam.ac.uk), which provide all well-established known associations of SNPs with their phenotypes, including subgenome-wide associations. We also used MR-PRESSO (MR Egger, Mendelian Randomization Pleiotropy RESidual Sum and Outlier) and multivariable MR to identify and correct for unknown potential pleiotropy. Using these genetic instruments, we validated that the effects on lipid profile were consistent with the well-established cholesterol-lowering effect of LA.[6]

Second, in the letter Lee makes a link between "limited numbers of IVs" and "bias from weak instruments"[1]; however, they are not equivalent. Instead, there is a "bias-variance trade-off for the number of instruments used in IV estimation".[7] Specifically, at a fixed mean F-statistic, increasing the number of instruments will lower the variance of the estimate (increase the precision) but at the same time may increase the possibility of bias from weak instruments.[7] The validity of the instrument is mainly based on the compliance with the MR assumptions rather than the number of instruments available. A single SNP, if validated, can also be used as an instrument in an MR study,[8] as has been the case in previous influential MR studies.[9 10] Lee did not provide any information about checking the instruments for associations with potential confounders, such as socioeconomic position, smoking and alcohol use, or checking for pleiotropic associations, in addition to sensitivity analysis using different analytic methods.[1]

We agree that using more valid instruments could increase the power of an MR study. However, we are unclear as to the validity of the use of 75 SNPs for LA as mentioned by Lee.[1] The 173 SNPs associated with LA at the genome-wide significance are highly correlated.[3] We cannot identify 75 independent SNPs meeting the selection criteria given by Lee ("linkage disequilibrium R^2 of 0.001, clumping distance of 10 000 kb, and a p-value threshold of 5.00E−08")[1]; those criteria only give the three SNPs providing the same information as what we used. However, if we apply a method suitable for correlated SNPs[11] and use all 167 SNPs available at genome-wide significance, we get an estimate very similar to that in our original letter (OR 0.97, 95% CI 0.95 to 0.98, p<0.001).

Jie V Zhao,[1] C Mary Schooling[1,2]

[1]School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China
[2]School of Public Health and Health Policy, City University of New York, New York City, New York, USA

Correspondence to Dr Jie V Zhao, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China; janezhao410@gmail.com

Handling editor Josef S Smolen

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent Not required.

Provenance and peer review Commissioned; internally peer reviewed.

To cite Zhao JV, Schooling CM. *Ann Rheum Dis* 2020;**79**:e29.

Received 5 December 2018
Accepted 5 December 2018
Published Online First 18 December 2018

Linked

► https://doi.org/10.1136/annrheumdis-2018-214810

Ann Rheum Dis 2020;**79**:e29. doi:10.1136/annrheumdis-2018-214830

REFERENCES

1 Lee YH. Role of linoleic acid in autoimmune disorders: a Mendelian randomisation study. *Ann Rheum Dis* 2020;79:e28.
2 Lawlor DA, Harbord RM, Sterne JA, *et al*. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* 2008;27:1133–63.
3 Guan W, Steffen BT, Lemaitre RN, *et al*. Genome-wide association study of plasma N6 polyunsaturated fatty acids within the cohorts for heart and aging research in genomic epidemiology consortium. *Circ Cardiovasc Genet* 2014;7:321–31.
4 May-Wilson S, Sud A, Law PJ, *et al*. Pro-inflammatory fatty acid profile and colorectal cancer risk: a Mendelian randomisation analysis. *Eur J Cancer* 2017;84:228–38.
5 Zhao JV, Schooling CM. Role of linoleic acid in autoimmune disorders: a Mendelian randomisation study. *Ann Rheum Dis* 2019;78:711–3.
6 Ramsden CE, Zamora D, Majchrzak-Hong S, *et al*. Re-evaluation of the traditional diet-heart hypothesis: analysis of recovered data from Minnesota Coronary Experiment (1968–73). *BMJ* 2016;353:i1246.
7 Burgess S, Thompson SG. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol* 2011;40:755–64.
8 Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res* 2017;26:2333–55.
9 Tchetgen Tchetgen EJ, Walter S, Glymour MM. Commentary: Building an evidence base for Mendelian randomization studies: assessing the validity and strength of proposed genetic instrumental variables. *Int J Epidemiol* 2013;42:328–31.
10 Au Yeung SL, Jiang C, Cheng KK, *et al*. Is aldehyde dehydrogenase 2 a credible genetic instrument for alcohol use in Mendelian randomization analysis in Southern Chinese men? *Int J Epidemiol* 2013;42:318–28.
11 Burgess S, Zuber V, Valdes-Marquez E, *et al*. Mendelian randomization with fine-mapped genetic data: choosing from large numbers of correlated instrumental variables. *Genet Epidemiol* 2017;41:714–25.

Ultrasound findings in palindromic rheumatism

We read the article by Mankia *et al* on the distinct ultrasound (US) imaging phenotype in palindromic rheumatism (PR) with great interest.[1] The authors found characteristic US findings in PR during flares that differ from those observed in patients with early rheumatoid arthritis (RA) or anti-cyclic citrullinated peptide (CCP)+ arthralgia. US extracapsular inflammation (periarticular inflammation, subcutaneous or peritendinous oedema), in most cases without joint synovitis, is the most frequent US finding in PR. These findings disappeared after acute attacks. The authors concluded that this imaging phenotype of extracapsular inflammation is specific for PR and may be distinguished from that observed in RA or persistent arthritis. They also suggest that true US intrasynovial inflammation may predict future RA in these patients.[1]

We analysed US findings in 54 patients with long-standing PR without evolution to RA or other forms of chronic arthritis in the intercritical phase of the disease.[2] Like Mankia *et al*,[1] we did not observe US synovitis in the intercritical period in most patients, even in anti-CCP+ patients, confirming the intermittent nature of this entity. However, US findings during the palindromic flare differed from those observed by Mankia *et al*. We made US evaluations during flares in 10 patients with PR and did not observe this characteristic US extracapsular image, although classical periarticular erythema was not found in any patient in the clinical examination. However, US active synovitis, defined as synovial hypertrophy ≥1 plus power Doppler signal, was found in six patients (60%).

We also analysed the rate of progression to RA or persistent arthritis in our cohort. In total, 13 out of 52 patients (2 patients were lost to follow-up) developed persistent arthritis (11 RA) after a mean follow-up of 4.8±1.6 years after study entry and US evaluation. Only two of the six patients with US synovitis during the disease flare evolved to RA in the follow-up. Furthermore, the percentage of patients who evolved to RA did not differ between patients with or without US synovitis in the intercritical period (31% vs 26%).

We have no satisfactory explanation for the discrepancies between the two studies, although patient characteristics differed, with a long-standing PR and most patients treated with disease-modifying antirheumatic drugs, mainly hydroxychloroquine, in our study, compared with a short PR disease duration in the study by Mankia *et al*.[1] The high median levels of C reactive protein (9.9 mg/dL) are quite surprising (could it be 9.9 mg/L?) as this is not a current finding in PR. In addition, differences in the sensitivity of the US equipment cannot be ruled out.

In conclusion, our findings in the intercritical phase of PR are similar to those observed by Mankia *et al* but those observed during the disease flare are not. We did not observe the typical isolated extracapsular inflammation (and it also seems less prevalent when the authors analysed the images using MRI) and US synovitis was not highly predictive of future RA. It is unclear whether these differences in the US pattern in the acute phase of PR may reflect differences in the type of PR populations (short duration vs long-standing disease) or drug therapies, and this merits further investigation.

Raimon Sanmarti,[1] Julio Ramírez,[1] Raul Castellanos-Moreira,[1] Sonia Raquel Cabrera-Villalba,[1] Virginia Ruiz-Esquide,[1] Georgina Salvador[2]

[1]Rheumatology Department, Hospital Clinic of Barcelona, Barcelona, Spain
[2]Rheumatology Department, Hospital Mútua de Terrassa, Barcelona, Spain

Correspondence to Dr Raimon Sanmarti, Department of Rheumatology, Hospital Clínic of Barcelona, Barcelona 08036, Spain; SANMARTI@clinic.ub.es

Handling editor Josef S Smolen

Contributors RS and SC-V designed the study. JR performed the ultrasound study. All the coauthors contributed to acquisition of data and manuscript writing.

Competing interests None declared.

Patient consent for publication Obtained.

Ethics approval Ethics Comitee of Hospital Clinic.

Provenance and peer review Not commissioned; internally peer reviewed.

To cite Sanmarti R, Ramírez J, Castellanos-Moreira R, *et al*. *Ann Rheum Dis* 2020;**79**:e30.

Received 27 November 2018
Accepted 2 December 2018
Published Online First 14 December 2018

Linked

► http://dx.doi.org/10.1136/annrheumdis-2018-214867

Ann Rheum Dis 2020;**79**:e30. doi:10.1136/annrheumdis-2018-214832

REFERENCES

1 Mankia K, D'Agostino MA, Wakefield RJ, *et al*. Identification of a distinct imaging phenotype may improve the management of palindromic rheumatism. *Ann Rheum Dis* 2018:annrheumdis-2018-214175.
2 Cabrera-Villalba S, Ramirez J, Salvador G, *et al*. Is there subclinical synovitis in patients with palindromic rheumatism in the intercritical period? a clinical and ultrasonographic study according to anticitrullinated protein antibody status. *J Rheumatol* 2014;41:1650–5.

Response to: 'Ultrasound findings in palindromic rheumatism' by Sanmarti *et al*

We thank Sanmarti *et al*[1] for their interest in our recent paper in which we describe the distinct imaging phenotype of palindromic rheumatism (PR).[2] As the authors point out, we identified a high prevalence of ultrasound (US) extra-capsular inflammation in flares of PR, often without coexistent US synovitis. This US pattern was specific for PR and may be useful in distinguishing PR from early persistent arthritis.

In an earlier US study of a Spanish PR cohort, Sanmarti and colleagues performed US assessment in 10 patients during flares of PR[3] and reported US power Doppler synovitis in 7, with 5 of these fulfilling criteria for US-defined synovitis. The authors did not identify US periarticular changes, and none of these patients had periarticular inflammation clinically. This imaging pattern is different to that identified in our cohort but, as the authors surmise, this could be explained by key differences in the respective PR cohorts. Our cohort consisted of patients with PR relatively early in their disease course (median 2.5 years), 90% of whom were naive for disease-modifying anti-rheumatic drugs (DMARDs). In contrast, patients with PR in the Spanish cohort had long-standing disease (median 11.6 years) and the majority were on treatment; 85% of patients had received DMARD therapy, with 61% established on at least one DMARD at the time of imaging. As such, it is possible that the imaging phenotype described in our study better reflects true de novo untreated PR and with more prolonged disease duration and/or therapy, this disease pattern may change. For example, it is possible that extra-capsular inflammation is an early phenomenon which may be suppressed by DMARD treatment. Indeed, US extra-capsular abnormalities (including tenosynovitis and/or periarticular soft tissue inflammation) without synovitis have also been described in a Chinese PR cohort, 62% of whom were DMARD-naive (the remaining 38% having received hydroxychloroquine).[4]

The complete absence of clinical periarticular inflammation during PR flares in the Spanish study is interesting and certainly differs from our experience. Indeed, periarticular inflammation was described as an important clinical hallmark both in the original description of PR[5] and subsequently.[6] The absence of this characteristic feature, perhaps due to the effect of therapy, is consistent with the imaging pattern observed by Sanmarti and colleagues.

We thank Sanmarti *et al* for identifying the error in units for median C reactive protein levels; this should be mg/L rather than mg/dL as suggested. While we compared the imaging pattern in PR with that seen in new-onset rheumatoid arthritis (RA) and anti-cyclic citrullinated peptide (CCP)-positive individuals with musculoskeletal symptoms (CCP+ at risk), we did not specifically address whether the flare imaging pattern is predictive of progression from PR to RA. This would require a larger longitudinal study and would certainly be an important area for future investigation. Interestingly although Sanmarti and colleagues report only two of their six patients with US synovitis during flare developed RA, Chen *et al* reported a significantly higher progression to RA in patients with US synovitis during flare compared with those without US synovitis (37.7% vs 3.7%, OR 15.05).[4] Intra-synovial power Doppler signal is also highly predictive of progression to clinical arthritis in CCP+ at risk, both at joint and patient level.[7]

Despite the recent studies on PR, there are many unanswered questions and the research agenda remains broad. The unique phenotype of this condition raises important questions about the pathogenesis and the optimal approach to treatment. Identifying biomarkers for accurate clinical risk prediction is also an important ambition. Well-phenotyped, treatment-naive inception cohorts will be crucial to furthering our understanding of this fascinating disease.

Kulveer Mankia,[1,2] Maria-Antonietta D'Agostino,[3] Paul Emery[1,2]

[1]Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Chapel Allerton Hospital, Leeds, UK
[2]NIHR Leeds Musculoskeletal Biomedical Research Centre, Leeds, UK
[3]Ambroise Paré Hospital, APHP, Université Versailles-Saint-Quentin en Yvelines - Rheumatology, Boulogne-Billancourt, France

Correspondence to Dr Kulveer Mankia, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Chapel Allerton Hospital, Leeds LS7 4SA, UK; k.s.mankia@leeds.ac.uk

Handling editor Josef S Smolen

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent Not required.

Provenance and peer review Commissioned; internally peer reviewed.

To cite Mankia K, D'Agostino M-A, Emery P. *Ann Rheum Dis* 2020;**79**:e31.

Received 11 December 2018
Accepted 12 December 2018
Published Online First 21 December 2018

Linked

► https://doi.org/10.1136/annrheumdis-2018-214832

Ann Rheum Dis 2020;**79**:e31. doi:10.1136/annrheumdis-2018-214867

REFERENCES

1 Sanmarti R, Ramírez J, Castellanos-Moreira R, *et al*. Ultrasound findings in palindromic rheumatism. *Ann Rheum Dis* 2020;79:e30.
2 Mankia K, D'Agostino MA, Wakefield RJ, *et al*. Identification of a distinct imaging phenotype may improve the management of palindromic rheumatism. *Ann Rheum Dis* 2018. 10.1136/annrheumdis-2018-214175. [Epub ahead of print 08 Oct 2018]
3 Cabrera-Villalba S, Ramirez J, Salvador G, *et al*. Is there subclinical synovitis in patients with palindromic rheumatism in the intercritical period? a clinical and ultrasonographic study according to anticitrullinated protein antibody status. *J Rheumatol* 2014;41:1650–5.
4 Chen H-H, Chen D-Y, Hsieh T-Y, *et al*. Predicting the progression of palindromic rheumatism to rheumatoid arthritis: the role of ultrasonography and anti-cyclic citrullinated peptide antibodies. *J Med Ultrasound* 2010;18:17–26.
5 Hench P, Rosenberg EF. Palindromic rheumatism. A 'new' oft recurring disease of joints (arthritis, periarthritis, para-arthritis) apparently producing no articular residues. Report of thirty-four cases; its relation to angio-neural arthrosis', 'allergic rheumatism' and rheumatoid arthritis. *Arch Int Med* 1944;73:293–321.
6 Guerne PA, Weisman MH. Palindromic rheumatism: part of or apart from the spectrum of rheumatoid arthritis. *Am J Med* 1992;93:451–60.
7 Nam JL, Hensor EM, Hunt L, *et al*. Ultrasound findings predict progression to inflammatory arthritis in anti-CCP antibody-positive patients without clinical synovitis. *Ann Rheum Dis* 2016;75:2060–7.

Can solid-phase assays replace immunofluorescence for ANA screening?

The paper by Pisetsky *et al*[1] has stimulated a timely and interesting debate on a focal point in the diagnosis of autoimmune rheumatic diseases, that is, the accuracy of the antinuclear antibody (ANA) test and the reliability of the results provided by this test. The data produced by Pisetsky *et al* once more demonstrate the poor standardisation of the ANA assay when performed by the indirect immunofluorescence (IIF) method and the enormous difference that exists between different ANA-IIF kits. However, issues are related to intermethods variability and to the intrinsic limitations of the IIF method. Whatever title is chosen, it entails either a relevant loss of diagnostic specificity or sensitivity. This raises the question of whether IIF on HEp-2 cell substrates should be still considered the *gold standard* for ANA detection as stated almost 10 years ago by the American College of Rheumatology (ACR). The reasons that led the ACR to take this position were related to the insufficient diagnostic sensitivity of emerging alternative methods to IIF. These alternative methods, which were then almost exclusively made up of immunoenzymatic assays (ELISA), were spreading in clinical laboratories as substitutes of the manual IIF method to overcome known IIF limitations and for their higher throughput. At that time, however, studies comparing the results of the IIF with the ELISA methods had shown that the ANA-IIF provided, in most cases, better performance than the ELISA methods,[2] despite IIF having low specificity and being non-sensitive in the detection of certain antibodies (such as Ro52, Ro60, ribosomes and Jo1) that play an important role in the diagnosis and classification of some ANA-associated rheumatic diseases (AARD) such as Sjögren's syndrome, systemic lupus erythematosus (SLE) and autoimmune inflammatory myopathies (AIM).

Ten years later, we must ask, 'Is this assumption still true?' As pointed out by Meroni *et al*[3] in their comment, in recent years, technology has made considerable progress. ELISA methods are now being abandoned and replaced by more accurate immunometric solid-phase assays (SPA) such as chemiluminescence and fluoroimmunoenzymatic methods. Over the years, these new SPAs have demonstrated high diagnostic performance, so the suggestion has been advanced that they could replace IIF for the detection of ANA, yielding a more objective, rapid and quantitative result.[4] Moreover, the consolidation of clinical laboratories, widely occurring both in Europe and in North America, has meant that the volume of autoantibody tests that each laboratory must perform today is greatly increased, making both manual and automated IIF techniques increasingly difficult to apply, because of their long turnaround time compared with the fully automated and random access SPA methods, and because the ANA-IIF test must be personally read and interpreted (under a microscope or video) and is therefore exposed to a certain degree of subjectivity and to the operator's experience.[5]

Two solid-phase monotest methods, called CTD screen, are available today for ANA screening which include a mixture of 15–16 purified native or recombinant antigens among those most frequently recognised by autoantibodies in AARD. We reviewed all the studies that have compared the diagnostic accuracy (sensitivity, specificity and efficiency) of the CTD screen assays versus IIF both in selected cohorts of patients with AARD and in the daily workup[6–12] (table 1), reproducing the real-life ANA testing as recommended in their correspondence by Infantino *et al*.[13]

Taken together, these studies show that IIF has a higher sensitivity and a much lower specificity than SPA. However, when these data are analysed using receiver operating characteristic curves and compared with an equal specificity value, even sensitivity is higher for SPA. Overall, these data suggest that screening by SPA yields results that are at least comparable to—and probably better than—ANA-IIF results. However, it is important to note that, with regard to the individual AARD, diagnostic accuracy is different. From the cited studies, it is evident that IIF is slightly superior to SPA in detecting SLE and scleroderma, while SPA methods guarantee better results in Sjögren's syndrome and in AIM. So, if today we should have to choose between one of the two methods, neither would allow us to diagnose all patients with AARD.

Therefore, from a clinical point of view, the best diagnostic strategy seems to be the combined use of the two methods, according to an algorithm which requires the subsequent identification of individual antibodies only in cases that are positive by SPAs. Would this strategy also be economically sustainable? A recent cost analysis has shown that screening by IIF followed by analysis of antibody fine specificity by immunometric or immunoblot methods in all ANA-IIF positive samples provides, in most cases, negative or clinically irrelevant results; and that, by using the two methods in parallel and proceeding with testing to identify the fine antibody specificity only in SPA positive samples, the costs associated with the many false ANA-IIF positives would be reduced,[12] avoiding in addition unnecessary clinical referrals and test repetitions.

A final issue regards the fact that the ANA-IIF test allows for identification of patterns such as mitochondrial, multiple nuclear dots, and rim-like, which have diagnostic significance for autoimmune liver diseases such as primary biliary cholangitis. The available evidence, therefore, suggests that SPAs are not yet able

Table 1 (A) Diagnostic accuracy and overall efficiency (correct classification rate) of the ANA-immunofluorescence (IIF) method compared with solid-phase assays (SPA). (B) The best performing method in ANA-associated autoimmune rheumatic diseases according to the different studies

A							B			
	Sensitivity (%)		Specificity (%)		Efficiency (%)					
Author/year	ANA-IIF	SPA	ANA-IIF	SPA	ANA-IIF	SPA	SLE	SjS	SSc	AIM
Op De Beeck *et al* (2011)[6]	87.2	73.0	86.3	96.9	86.6	88.6	SPA	SPA	IIF	SPA
Robier *et al* (2016)[7]	98.7	82.7	85.8	98.2	81.1	90.7	IIF	SPA	IIF	–
Bentow *et al* (2015)[9]*	84.8	78.1	64.7	94.1	81.0	86.6	IIF	SPA	SPA	–
Otten *et al* (2017)[8]	81.7	78.9	88.6	95.1	72.8	77.1	IIF	SPA	Equal	IIF
van der Pol *et al* (2018)[10]†	90.0	95.1	76.0	80.0	78.9	83.4	SPA	SPA	Equal	Equal
Claessens *et al* (2018)[11]†*	95.2	83.1	61.0	92.9	69.8	89.1	IIF	SPA	Equal	SPA
Bizzaro *et al* (2018)[12]†	89.2	87.1	64.6	98.0	69.9	96.0	IIF	SPA	IIF	SPA

*In this study, ANA-IIF reading was performed with a computer-aided system (Nova View, Inova Diagnostics).
†These studies compared IIF with two SPA methods (chemiluminescence and fluoroimmunoenzymatic assay); data for SPA methods are combined.
AIM, autoimmune inflammatory myopathies; ANA, antinuclear antibody; IIF, indirect immunofluorescence; SLE, systemic lupus erythematosus; SSc, systemic sclerosis; SjS, Sjögren's syndrome.

to completely replace IIF and that the IIF method should be used for ANA screening until solid-phase methods become available to detect a greater number of autoantibodies not yet present in the antigenic panel of CTD screen assays. Only then will it be possible to evaluate whether these new methods would actually be able to completely replace the ANA-IIF method.

Nicola Bizzaro

Laboratory of Clinical Pathology, San Antonio Hospital, Tolmezzo–Azienda Sanitaria Universitaria Integrata di Udine, Tolmezzo, Italy

Correspondence to Dr Nicola Bizzaro , Laboratory of Clinical Pathology, San Antonio Hospital, Tolmezzo–Azienda Sanitaria Universitaria Integrata di Udine, Tolmezzo 33028, Italy ; nic.bizzaro@gmail.com

Handling editor Josef S Smolen

Contributors NB is the sole author of the article.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests The author has received speaker honoraria from Thermo Fisher and Werfen.

Patient consent Not required.

Provenance and peer review Not commissioned; internally peer reviewed.

To cite Bizzaro N. *Ann Rheum Dis* 2020;**79**:e32.

Received 22 November 2018
Accepted 25 November 2018
Published Online First 8 December 2018

Linked

► http://dx.doi.org/10.1136/annrheumdis-2018-214829

Ann Rheum Dis 2020;**79**:e32. doi:10.1136/annrheumdis-2018-214805

REFERENCES

1 Pisetsky DS, Spencer DM, Lipsky PE. Assay variation in the detection of antinuclear antibodies in the sera of patients with established SLE. *Ann Rheum Ds* 2018;77:911–3.
2 Tonuttia E, Bassetti D, Piazza A, *et al*. Diagnostic accuracy of ELISA methods as an alternative screening test to indirect immunofluorescence for the detection of antinuclear antibodies. Evaluation of five commercial kits. *Autoimmunity* 2004;37:171–6.
3 Meroni PL, Chan EK, Damoiseaux J, *et al*. Unending story of the indirect immunofluorescence assay on HEp-2 cells: old problems and new solutions? *Ann Rheum Dis* 2019;78:e46.
4 Bossuyt X, Fieuws S. Detection of antinuclear antibodies: added value of solid phase assay? *Ann Rheum Dis* 2014;73:e10.
5 Mahler M. Lack of standardisation of ANA and implications for drug development and precision medicine. *Ann Rheum Dis* 2019;78:e33.
6 Op De Beeck K, Vermeersch P, Verschueren P, *et al*. Detection of antinuclear antibodies by indirect immunofluorescence and by solid phase assay. *Autoimmun Rev* 2011;10:801–8.
7 Robier C, Amouzadeh-Ghadikolai O, Stettin M, *et al*. Comparison of the clinical utility of the Elia CTD Screen to indirect immunofluorescence on HEp2 cells. *Clin Chem Lab Med* 2016;54:1365–70.
8 Otten HG, Brummelhuis WJ, Fritsch-Stork R, *et al*. Measurement of antinuclear antibodies and their fine specificity: time for a change in strategy? *Clin Exp Rheumatol* 2017;35:462–70.
9 Bentow C, Lakos G, Rosenblum R, *et al*. Clinical performance evaluation of a novel, automated chemiluminescent immunoassay, QUANTA Flash CTD Screen Plus. *Immunol Res* 2015;61:110–6.
10 van der Pol P, Bakker-Jonges LE, Kuijpers JHSAM, *et al*. Analytical and clinical comparison of two fully automated immunoassay systems for the detection of autoantibodies to extractable nuclear antigens. *Clinica Chimica Acta* 2018;476:154–9.
11 Claessens J, Belmondo T, De Langhe E, *et al*. Solid phase assays versus automated indirect immunofluorescence for detection of antinuclear antibodies. *Autoimmun Rev* 2018;17:533–40.
12 Bizzaro N, Brusca I, Previtali G, *et al*. The association of solid-phase assays to immunofluorescence increases the diagnostic accuracy for ANA screening in patients with autoimmune rheumatic diseases. *Autoimmun Rev* 2018;17:541–7.
13 Infantino M, Manfredi M, Soda P, *et al*. ANA testing in 'real life'. *Ann Rheum Dis* 2020;79:e3.

Response to: 'Can solid-phase assays replace immunofluorescence for ANA screening?' by Bizzaro

We would like to thank Dr Bizzaro for his commentary[1] on our article[2] on the variability of testing for antinuclear antibodies (ANA) by indirect immunofluorescence (IIF). Along with other letters that have been published in response to our article,[3–9] Dr Bizzaro's letter highlights the concerns about the IIF, its status as the 'gold standard' and the availability of other technologies (eg, solid phase assays) that alone or together can provide testing with comparable or better sensitivity and specificity than the IIF. As Dr Bizzaro indicates, the utilisation of these technologies may have advantages in terms of overall costs of patient care.

We agree with Dr Bizzaro that the role of different assay approaches must be evaluated and interpreted in the context of the clinical setting and that the issues for screening may differ for classes of diseases (eg, connective tissue disease and autoimmune liver disease) as well as individual diseases (eg, systemic lupus erythematosus and Sjogren's syndrome). As we have discussed, for systemic lupus erythematous, the stakes for testing are high since ANA positivity is used as a criterion for disease classification, entry into clinical trials and prescription of medications for products approved for active, autoantibody positive disease.

At this point, we think that it is time for professional organisations and regulatory agencies to recognise the strong evidence for assay variability and start the process of evaluating different assays and platforms for specific purposes and provide guidance for better standardisation. An important first step may be to reopen the question of whether there is in fact a 'gold standard' for ANA testing in general and then determine the best test(s) for specific applications. We are glad that our article has sparked so many letters and believe that the data and ideas presented indicate clearly that re-evaluation of ANA is essential in view of new technologies and new uses for this venerable and widely performed test.

David S Pisetsky,[1] Diane M Spencer,[1] Peter E Lipsky,[2] Brad H Rovin[3]

[1]Department of Medicine and Immunology, Duke University Medical Center and Medical Research Service, VA Medical Center, Durham, North Carolina, USA
[2]RILITE Research Institute, Charlottesville, Virginia, USA
[3]Division of Nephrology, The Ohio State University, Wexner Medical Center, Columbus, Ohio, USA

Correspondence to Dr David S Pisetsky, Department of Medicine and Immunology, Duke University Medical Center and Medical Research Service, VA Medical Center, Durham, NC 27705, USA; david.pisetsky@duke.edu

Handling editor Josef S Smolen

Competing interests None declared.

Patient consent Not required.

Provenance and peer review Commissioned; internally peer reviewed.

To cite Pisetsky DS, Spencer DM, Lipsky PE, *et al*. *Ann Rheum Dis* 2020;**79**:e33.

Received 2 January 2019
Accepted 3 January 2019
Published Online First 18 January 2019

Linked

► http://dx.doi.org/10.1136/annrheumdis-2018-214805

Ann Rheum Dis 2020;**79**:e33. doi:10.1136/annrheumdis-2018-214829

REFERENCES

1 Bizzaro N. Can solid-phase assays replace immunofluorescence for ANA screening? *Ann Rheum Dis* 2020;79:e32.
2 Pisetsky DS, Spencer DM, Lipsky PE, *et al*. Assay variation in the detection of antinuclear antibodies in the sera of patients with established SLE. *Ann Rheum Dis* 2018;77:911–3.
3 Infantino M, Manfredi M, Soda P, *et al*. ANA testing in 'real life'. *Ann Rheum Dis* 2020;79:e3.
4 Mahler M. Lack of standardisation of ANA and implications for drug development and precision medicine. *Ann Rheum Dis* 2019;78:e33.
5 Meroni PL, Chan EK, Damoiseaux J, *et al*. Unending story of the indirect immunofluorescence assay on HEp-2 cells: old problems and new solutions? *Ann Rheum Dis* 2019;78:e46.
6 Pacheco Y, Monsalve DM, Acosta-Ampudia Y, *et al*. Antinuclear autoantibodies: discordance among four different assays. *Ann Rheum Dis* 2020;79:e6.
7 Pregnolato F, Borghi MO, Meroni PL, *et al*. Pitfalls of antinuclear antibody detection in systemic lupus erythematosus: the positive experience of a national multicentre study. *Ann Rheum Dis* 2019;78:e50.
8 Van Hoovels L, Schouwers S, Van den Bremt S, *et al*. Variation in antinuclear antibody detection by automated indirect immunofluorescence analysis. *Ann Rheum Dis* 2019;78:e48.
9 Willems P, De Langhe E, Westhovens R, *et al*. Antinuclear antibody as entry criterion for classification of systemic lupus erythematosus: pitfalls and opportunities. *Ann Rheum Dis* 2019;78:e76.

Is it time to redefine the role of low-dose radiotherapy for benign disease?

It is estimated that 24% of the general adult population is currently suffering from osteoarthritis (OA), affecting 10% of men and 18% of women over 60 years of age in high-income countries. A WHO report predicted that degenerative OA will become the fourth leading cause of disability by 2020.[1] This may not only affect the individuals who suffer from the diseases, but will undeniably have an impact on national health systems in social and economic terms.

There is no specific or definitive treatment for the early and late stages of degenerative OA. Weight loss, maintaining moderate levels of exercise and physical rehabilitation approaches (local heat, magnetic therapy and shock waves, among others) are some of the conservative therapies applied. Analgesics and non-steroidal anti-inflammatory drugs, symptomatic slow-acting drugs for OA, corticosteroids, anaesthetics and other local injections have been proposed for the relief of the symptoms before a prosthetic replacement of the damaged joint would finally be carried out at the end of a long road. None of these options have demonstrated high efficacy, and even more importantly can provoke multiple side effects and acute and late morbidities (ie, gastrointestinal bleeding, kidney and cardiac disorders, and so on) which may become serious and even compromise the patient's life.[2]

The clinical effectiveness of low-dose radiation therapy (LD-RT) in the range of 0.3–0.7 Gy single dose and 3–10 Gy total dose for pain relief and subsequent improvement of joint functionality has been recognised for several decades. Further, the anti-inflammatory efficacy of LD-RT has been confirmed in several experimental models, both in vitro and in vivo.[3–7] The first clinical evidence of its efficacy in non-cancerous osteoarticular disorders dates from the end of the 19th century, although there has traditionally been some resistance for its widespread use due to the fear of its possible side effects and carcinogenesis. However, the clinical experience of using LD-RT acquired in recent years regarding its radiobiological and immunological mechanisms of action,[8 9] its low toxicity profile and its proven effectiveness in degenerative OA has reinforced its role as a therapeutic alternative in these patients without other options. This has been evidenced by a multitude of trials. In addition, radiotherapy is a non-invasive treatment that does not interfere with other therapies, something of great importance considering most candidate patients' multimorbidity characteristics.

In 2017, at the 37th European SocieTy fo Radiotherapy and Oncology (ESTRO) Annual Congress, Minten *et al*[10–12] presented and has since published the results of two double-blinded randomised trials on the effect of LD-RT therapy for symptomatic relief and functional improvement of degenerative OA of the hand or knee joints. They provided the first clinical studies that compared a modern radiation therapy technique with a sham irradiated group with identical patient and disease characteristics. In the first study, the authors analysed the results observed in 56 patients with OA of the hand (finger joint OA or rhizarthrosis), while in the second study 55 patients with OA of the knee were enrolled, applying the same randomised, double-blinded design of radiotherapy at low dose (6 Gy in fractions of 1 Gy, 3 fractions/week) versus sham radiotherapy. In both studies, the authors evaluated the clinical response at 3 months of treatment according to the Outcome Measures in Rheumatology-Osteoarthritis Research Society International (OMERACT-OARSI) response criteria, including evaluation of pain and functionality of the treated joints. At 3 months' follow-up, the authors did not observe any differences in any of the response parameters analysed between the group of patients who received radiotherapy and those who did not (sham treatment).

The results of these two well-designed studies raise some questions regarding the effectiveness of radiotherapy at low doses in degenerative osteoarticular disorders from the perspective of radiation oncologists with several years of clinical experience in the use of LD-RT for the symptomatic relief of these diseases. Several aspects should be taken into account when definitively evaluating the possible negative impact of these findings.

First, under the term 'osteoarticular pathology', very different entities are included, covering both OA and enthesopathies. Although radiotherapy is effective in the symptomatic treatment of osteoarticular disorders, it is well described that a higher rate of complete pain remissions is achieved in the treatment of patients with calcaneodynia, achillodynia, bursitis trochanterica and shoulder syndrome (enthesopathies) than in the treatment of gonarthrosis. Degenerative OA of the knee or of the interphalangeal joints of the hand is a chronic disorder with destruction of the bone and cartilage; although radiotherapy can alleviate the inflammation and pain symptoms secondary to the joint destruction, the underlying pathophysiological mechanisms will continue without evident changes. Therefore, the analgesic effect is lower than that observed in other disorders.[13 14]

Second, since 2000, at least six clinical studies have been published on the efficacy of LD-RT in degenerative knee or hand OA, including a total of 1508 patients who were analysed retrospectively[13 15–17] or prospectively[14 18]. Irradiation doses ranged from 3 to 6 Gy total dose with a fractionation of 0.5–1 Gy single dose and 2–3 fractions per week. With a median follow-up of 3–48 months (median 29 months), the response rate mainly referring to pain relief ranged from 63% to 90% in the different clinical series. However, between 7% and 100% of the patients (median 15%) required a second course of radiotherapy 6–12 weeks later to reach a positive clinical outcome if the initial result had not been completely satisfactory. Further, Mücke *et al*[19] collected data from 238 institutions in Germany, of which 188 (79%) used LD-RT for the treatment of knee OA. The authors reviewed data from 4544 patients treated in 2008 with a median dose of 6 Gy (range 3–12 Gy) by two or three weekly fractions of 1 Gy of median dose (range 0.25–3 Gy). Thirty per cent of patients received a second series of radiotherapy 6–12 weeks after completion of the first. The authors observed symptomatic pain relief in 79.5% of patients.[19] Thus, we believe that these studies clearly show the real achievable goals that LD-RT can produce.

Third, the striking question raised by the two articles is whether LD-RT has to be definitively dropped from our protocols for OA, as suggested by the authors, or it has to be restricted only to the most refractory patients. Usually LD-RT for OA is delivered to patients with very chronic disease and is unsuitable for other treatments. Thus, the assumed goal of improvement in 40% of the patients in the LD-RT arm is considered too highly optimistic, being a drawback of the studies. Another important aspect is that the small number of patients could raise doubts regarding the statistical results. The percentage of patients who responded to the placebo was unexpectedly high. Further, the inclusion of patients with a higher body mass index in the LD-RT group might falsify the results. It has become obvious that overweight persons do have a permanent higher basal level of inflammation (summarised in ref 20) and a direct comparison with the placebo group is therefore difficult.

Furthermore, a delayed onset of the analgesic effects of LD-RT was established previously and results showed a significant improvement in long-term efficacy compared with results

obtained immediately after radiotherapy.[8 9 13] In the two randomised studies, however, the evaluation of the outcome was limited to 3 months after completion of treatment. Thus, no long-term benefit was evaluated at 6 or 12 months. In addition, the studies did not offer to carry out a second series of irradiation, which, according to the experience of previous clinical studies, can benefit a high percentage of patients.[13–19]

The design and evaluation of both randomised trials, including different clinical questionnaires and the assessment of quality of life, are good, but the additional biochemical inflammatory parameters (erythrocyte sedimentation rate and C reactive protein serum levels) do not appear to be very useful criteria for assessing inflammatory response in chronic arthrodegenerative disease. Against this background, Rühle *et al*[9] recently reported on a modulation of T cells and monocytes and a reduction of the activation marker CD69 on T, B and NK cells in the blood of patients with chronic painful musculoskeletal diseases following radon spa treatment.[19] Comparable assays are currently running in patients with LD-RT (NCT02653079) and complementary ones are planned in Spain in the near future.

Additionally, it is important to further take into consideration the very long clinical history and treatment prior to the application of LD-RT, while both studies restricted the inclusion criteria to the duration of symptoms of more than or equal to 5 years in 68% in the LD-RT group and 54% in the sham-treated group.

Finally, the clinical data provided on irradiation volumes, at least with regard to the treatment of arthrosis of the thumb, raise concerns about their suitability, given that a certain relationship had been established previously between the size of the field and response to treatment, with larger fields than those used by Minten *et al* associated with a higher response to treatment, and that to some extent might contribute to the low rate of responses observed in these patients.[17]

In conclusion, the two studies raise very interesting questions, and their extraordinarily accurate design should serve as a basis for future clinical studies that contemplate on the efficacy of LD-RT, and not only restricted to the hand and knee joints. Moreover, the adequate definition of volumes of irradiation, inclusion of a second series of treatment and the evaluation of a long-term response beyond 3 months after the treatment might further contribute to a more accurate selection of patients that most probably will benefit from LD-RT. The future work from these studies is to define the patients who are prone to clinical improvement after LD-RT and to develop biomarkers to predict responses to LD-RT.

Angel Montero,[1] Sebastia Sabater,[2] Franz Rödel,[3] Udo S Gaipl,[4] Oliver J Ott,[4] Michael Heinrich Seegenschmiedt,[5] Meritxell Arenas[6]

[1]Department of Radiation Oncology, Hospital Universitario HM Sanchinarro, Madrid, Spain
[2]Department of Radiation Oncology, Complejo Hospitalario Universitario de Albacete, Albacete, Spain
[3]Department of Radiotherapy and Oncology, Universitätsklinikum Frankfurt am Main, Frankfurt am Main, Germany
[4]Department of Radiation Oncology, Universitätsklinikum Erlangen, Erlangen, Germany
[5]Department of Radiation Oncology, Strahlenzentrum Hamburg, Hamburg, Germany
[6]Department of Radiation Oncology, Hospital Universitari Sant Joan de Reus, University of Rovira i Virgili, Tarragona, Spain

Correspondence to Dr Meritxell Arenas, Department of Radiation Oncology, Hospital Universitari Sant Joan de Reus, University of Rovira i Virgili, Tarragona, Spain; meritxell.arenas@gmail.com

Handling editor Josef S Smolen

Contributors All the authors wrote and reviewed the manuscript.

Competing interests None declared.

Patient consent Not required.

Provenance and peer review Not commissioned; internally peer reviewed.

To cite Montero A, Sabater S, Rödel F, *et al*. *Ann Rheum Dis* 2020;**79**:e34.

Received 4 December 2018
Accepted 8 December 2018
Published Online First 21 December 2018

Linked

► http://dx.doi.org/10.1136/annrheumdis-2018-214896

Ann Rheum Dis 2020;**79**:e34. doi:10.1136/annrheumdis-2018-214873

REFERENCES

1 Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. *Bull World Health Organ* 2003;81:646–56.
2 Wieland HA, Michaelis M, Kirschbaum BJ, *et al*. Osteoarthritis - an untreatable disease? *Nat Rev Drug Discov* 2005;4:331–44.
3 Rödel F, Keilholz L, Herrmann M, *et al*. Radiobiological mechanisms in inflammatory diseases of low-dose radiation therapy. *Int J Radiat Biol* 2007;83:357–66.
4 Arenas M, Sabater S, Hernández V, *et al*. Anti-inflammatory effects of low-dose radiotherapy. *Strahlentherapie und Onkologie* 2012;188:975–81.
5 Arenas M, Gil F, Gironella M, *et al*. Anti-inflammatory effects of low-dose radiotherapy in an experimental model of systemic inflammation in mice. *Int J Radiat Oncol Biol Phys* 2006;66:560–7.
6 Arenas M, Gil F, Gironella M, *et al*. Time course of anti-inflammatory effect of low-dose radiotherapy: correlation with TGF-beta(1) expression. *Radiother Oncol* 2008;86:399–406.
7 Deloch L, Derer A, Hueber AJ, *et al*. Low-dose radiotherapy ameliorates advanced arthritis in hTNF-α tg mice by particularly positively impacting on bone metabolism. *Front Immunol* 2018;9:1834.
8 Rödel F, Frey B, Manda K, *et al*. Immunomodulatory properties and molecular effects in inflammatory diseases of low-dose x-irradiation. *Front Oncol* 2012;2:120.
9 Rühle PF, Wunderlich R, Deloch L, *et al*. Modulation of the peripheral immune system after low-dose radon spa therapy: Detailed longitudinal immune monitoring of patients within the RAD-ON01 study. *Autoimmunity* 2017;50:133–40.
10 Minten MJM, Mahler EAM, Leseman-Hoogenboom MM, *et al*. PV-0468: Low-dose radiation therapy as treatment for hand and knee osteoarthritis: two double-blinded RCTs. *Radiotherapy and Oncology* 2018;127:S240.
11 Minten MJM, Leseman-Hoogenboom MM, Kloppenburg M, *et al*. Lack of beneficial effects of low-dose radiation therapy on hand osteoarthritis symptoms and inflammation: a randomised, blinded, sham-controlled trial. *Osteoarthritis Cartilage* 2018;26:1283–90.
12 Mahler EAM, Minten MJ, Leseman-Hoogenboom MM, *et al*. Effectiveness of low-dose radiation therapy on symptoms in patients with knee osteoarthritis: a randomised, double-blinded, sham-controlled trial. *Ann Rheum Dis* 2018:annrheumdis-2018-214104.
13 Keller S, Müller K, Kortmann RD, *et al*. Efficacy of low-dose radiotherapy in painful gonarthritis: experiences from a retrospective East German bicenter study. *Radiat Oncol* 2013;8:29.
14 Micke O, Ugrak E, Bartmann S, *et al*. Radiotherapy for calcaneodynia, achillodynia, painful gonarthrosis, bursitis trochanterica, and painful shoulder syndrome - Early and late results of a prospective clinical quality assessment. *Radiat Oncol* 2018;13:71.
15 Glatzel M, Frohlich D, Kraub A. Results of radiotherapy for. gonarthrosis. *Benign News* 20023:9–11.
16 Ruppert R, Seegenschmiedt MH, Sauer R. [Radiotherapy of osteoarthritis. Indication, technique and clinical results]. *Orthopade* 2004;33:56–62.
17 Kaltenborn A, Bulling E, Nitsche M, *et al*. The field size matters: low dose external beam radiotherapy for thumb carpometacarpal osteoarthritis : Importance of field size. *Strahlenther Onkol* 2016;192:582–8.
18 Micke O, Seegenschmiedt MH, Adamietz IA, *et al*. Low-Dose Radiation Therapy for Benign Painful Skeletal Disorders: The Typical Treatment for the Elderly Patient? *Int J Radiat Oncol Biol Phys* 2017;98:958–63.
19 Mücke R, Seegenschmiedt MH, Heyd R, *et al*. [Radiotherapy in painful gonarthrosis. Results of a national patterns-of-care study]. *Strahlenther Onkol* 2010;186:7–17.
20 Neumann E, Junker S, Schett G, *et al*. Adipokines in bone disease. *Nat Rev Rheumatol* 2016;12:296–302.

Response to: 'Is it time to redefine the role of low-dose radiotherapy for benign disease?' by Montero *et al*

We thank the authors, with affiliations of six departments of radiation oncology in Spain and Germany, for their interest in our randomised controlled trials (RCT) and their compliments on the study designs.[1] Our studies showed no substantial beneficial effect on symptoms of low-dose radiation therapy (LDRT) in patients with knee and hand osteoarthritis (OA).[2 3] These findings dispute the effectiveness of radiation therapy at low doses in OA as commonly used in certain parts of the world. However, the authors mention several aspects, which should be taken into account when definitively appreciating our results, which we like to reply on in this letter.

We agree that OA is a serious health problem considering its high clinical burden, the high and rising prevalence, and the growing impact on healthcare and future economic costs.[4] Subsequently, there is a clear need to improve management of OA that is supported by scientific evidence, as no effective disease-modifying treatments are available. Therefore, current treatment focuses primarily on the reduction of symptoms including pain and loss of function while the importance and efficacy of non-surgical treatment modalities have been described in several international clinical guidelines for the management of knee and hand OA.[5 6] Of note, the authors mention approaches (ie, local heat, magnetic therapy and shock wave) not supported by evidence or included in international guidelines for the management of OA.

The authors state that the analgesic effect of radiation therapy is smaller in OA than other osteoarticular disorders (eg, calcaneodynia, achillodynia, bursitis trochanterica). This statement is based on findings of two observational studies with heterogeneous populations using a single transition question (von Pannewitz scale) that is likely to be biased by social desirability, in particular when assessed by telephone. The inferior design of those studies does not allow evidence-based discrimination in effects between different patient groups. Therefore, well-designed randomised studies in well-defined patient groups are necessary.

Remarkably, the authors mention that the clinical effectiveness of LDRT has been recognised for several decades and that the clinical effectiveness is proven by a multitude of trials. However, the authors ignore the results of our systematic literature review summarising the results of seven clinical observational studies.[7] Indeed, high improvement rates were reported in those studies. However, the methodological quality of all studies was judged as weak (no blinding, retrospective designs, uncontrolled studies and non-validated single-item outcome measures). Therefore, we concluded that there is insufficient high-level evidence available to indisputably demonstrate the effectiveness of LDRT in patients with OA. In addition, two low-quality RCTs in patients with OA were published in the 1970s and showed no effect of a higher dose radiation therapy than recommended in current guidelines.[7] Thus, in our opinion there is insufficient evidence to justify the use of LDRT for OA in clinical practice, which was exactly the motivation for the setting up of our trials.

The authors question several methodological aspects of our study, that is, the assumed and found placebo effect (40% response in the sham group), the unbalance in body mass index (BMI) between groups, the timing of the primary endpoint (3 months after intervention), the validity of the treatment protocol, length of follow-up and lack of assessment of biochemical inflammatory parameters other than erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP). We will address these comments point by point:

- We based our assumption of a high placebo effect on previous research reporting on the power of placebo for pain relief in OA, in particular for rather invasive interventions such as sham LDRT, which are associated with higher placebo effects.[8] The 3-month response in knee OA was about 40% in both groups, confirming our assumption and illustrating the substantial effect of placebo and regression to the mean. In addition, LDRT can only have a place in clinical practice when its effect would outweigh the time investment, patients' burden, radiation exposure and costs. It is very likely that placebo and regression to the mean effects are also responsible for the reported improvements of LDRT on symptoms in previous studies suffering from methodological shortcomings.
- The authors assume that a higher BMI in the LDRT group (knee OA) could have affected our results because overweighed persons have a higher level of inflammation. We agree on this point as we cannot rule out BMI as potential confounder due to potential unbalanced randomisation considering the limited sample size. However, as median BMI (and thus inflammation level) was higher in the LDRT group than in the sham group, a potential overestimation of effect is more likely than an underestimation as regression to the mean effect is more likely in the LDRT group. As described in the results section, additional analyses adjusting for BMI did not modify our results. In summary, we do not have any reasons to assume that differences between groups in BMI might jeopardise the results.
- The authors challenge our choice of primary endpoint at 3 months. We hypothesised that short-term effects of LDRT on inflammation (and thus on pain) are more likely than long-term effects. Moreover, we followed the 2015 guidelines for radiation therapy of benign diseases of the German Society for Radio-oncology (DEGRO) stating that evaluation of the treatment effect should be performed after 2–3 months. Nevertheless, clinical results at 6 and 12 months (manuscript in preparation) show invariably no relevant differences between groups in both knee and hand OA trials after long-term follow-up.
- Furthermore, the authors state that low response rates could be attributed to the limited size of the radiation fields. We followed the 2015 DEGRO guidelines recommending that target volumes should include joint cartilage adjoining bony structures, synovial tissue, and adjoining muscles and connective tissues. The total dose should range from 3.0 to 6.0 Gy, with fraction sizes of 0.5–1.0 Gy, applied two to three times per week. All these recommendations were followed in both our studies. A new RCT would be necessary to examine the hypothesis of the authors that a larger field or a smaller fraction size would have resulted in a more positive effect.
- Other than suggested by the authors we did not exclude patients on the basis of symptom duration. In line with the DEGRO guidelines we included the relevant patients, being patients who failed to respond to conservative treatment. Nevertheless, additional analysis with adjustment for symptom duration as potential confounder yielded similar results.
- The authors suggest to use biochemical inflammatory parameters other than ESR and CRP (eg, T and B cells, monocytes)

to assess the inflammatory response. Indeed, previous in vitro and in vivo studies of OA in animal models have shown that LDRT exerts anti-inflammatory effects.[9] However, in humans there is currently no high-level evidence that supports the hypothesis that an anti-inflammatory response leads towards substantial reduction of symptoms in OA. Of note, we did not observe a substantial reduction of symptoms or inflammatory signs. We plan to assess the effect of LDRT on the proinflammatory protein S100A8/A9. To our knowledge, the use of monocytes and T cells to assess inflammatory response in OA is not yet generally accepted and the results reported by Rühle *et al* did not confirm the usefulness of those parameters to assess inflammation in OA.[10] They only observed small fluctuating changes in some parameters during a period of 30 weeks in a heterogeneous sample without data on the clinical diagnosis.

Recently, the evidence for the effectiveness of LDRT for benign (musculoskeletal) diseases has been reviewed.[11] McKeown *et al* conclude that in the UK the use of radiation therapy for benign conditions is limited, in contrast to practice in Germany.[11] They also conclude that interpretation of the literature on radiation therapy for benign conditions is problematic because much of the evidence is based on case reports and single institution case series, although some randomised studies and systematic reviews do exist. There is a need to question and discuss the necessity of treatments commonly used but not supported by evidence. In recent years, this problem has gained more attention and the internationally expanding Choosing Wisely campaign is a good example of the effort taken to decrease tests and treatments that do not have additional value for patients and may even cause harm.[12] We therefore recommend to add LDRT treatment for other benign (musculoskeletal) disorders to the Choosing Wisely list of the European Society for Radiotherapy and Oncology.

Finally, when taking a more reflective and contemplating position, it should be noted that the scenario that unfolds here is not uncommon in history of medicine. There are numerous examples of treatments that have been used for decades and were considered beneficial, based on uncontrolled studies, until higher quality evidence demonstrated that the treatment was not effective. Well-known examples include, for example, hormone replacement therapy for cardiovascular disease in postmenopausal women,[13] steroids after head trauma[14] and surgery in lumbosacral radicular syndrome in the acute phase of the disease.[15] The first results regarding those treatments from high-quality trials were received by disbelief, much alike the current situation. However, arguments that the clinical effectiveness of LDRT has been recognised for several decades and that results of observational studies are positive are not valid. The way forward is clear: the burden of proof to demonstrate effectiveness of LDRT in OA lies with its proponents, and until then, use of this treatment should not be advocated.

In conclusion, considering the consistency of findings of both our trials and the lack of high-level evidence showing the opposite, we feel that it is time indeed to redefine the role of LDRT in knee and hand OA and that deimplementation of LDRT in clinical practice should be seriously and urgently considered. However, we acknowledge the importance of replication to further strengthen the body of knowledge by conducting preregistered well-designed randomised trials with validated outcomes.

Elien A M Mahler ,[1] **Michiel JM Minten**,[1] **Mathilde M Leseman-Hoogenboom**,[2] **Philip M P Poortmans**,[2,3] **Jan Willem Leer**,[2] **Simone S Boks**,[4] **Frank H J van den Hoogen**,[5] **Alfons A den Broeder**,[1] **Cornelia H van den Ende**[1]

[1]Department of Rheumatology, Sint Maartenskliniek Nijmegen, Nijmegen, The Netherlands
[2]Department of Radiation Oncology, Radboud University Medical Center, Nijmegen, The Netherlands
[3]Department of Radiation Oncology, Institut Curie, Paris, France
[4]Department of Radiology, Sint Maartenskliniek Nijmegen, Nijmegen, The Netherlands
[5]Department of Rheumatology, Radboud University Medical Center, Nijmegen, The Netherlands

Correspondence to Elien A M Mahler, Department of Rheumatology, Sint Maartenskliniek, Nijmegen 6500 GM, The Netherlands; e.mahler@maartenskliniek.nl

Handling editor Josef S Smolen

Contributors Drafting of the manuscript: EAMM, CHMvdE. Critical revision of the manuscript for important intellectual content: all authors. Final approval of the manuscript: all authors. All authors take responsibility for the integrity of the work and agreed to submit the response for publication.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Commissioned; internally peer reviewed.

To cite Mahler EAM, Minten MJM, Leseman-Hoogenboom MM, *et al*. *Ann Rheum Dis* 2020;**79**:e35.

Received 13 January 2019
Accepted 20 January 2019
Published Online First 1 February 2019

Linked

► http://dx.doi.org/10.1136/annrheumdis-2018-214873

Ann Rheum Dis 2020;**79**:e35. doi:10.1136/annrheumdis-2018-214896

ORCID iD
Elien A M Mahler http://orcid.org/0000-0002-6833-3157

REFERENCES

1 Montero A, Sabater S, Rödel F, *et al*. Is it time to redefine the role of low-dose radiotherapy for benign disease? *Ann Rheum Dis* 2020;79:e34.
2 Mahler EAM, Minten MJ, Leseman-Hoogenboom MM, *et al*. Effectiveness of low-dose radiation therapy on symptoms in patients with knee osteoarthritis: a randomised, double-blinded, sham-controlled trial. *Ann Rheum Dis* 2019;78:83–90.
3 Minten MJM, Leseman-Hoogenboom MM, Kloppenburg M, *et al*. Lack of beneficial effects of low-dose radiation therapy on hand osteoarthritis symptoms and inflammation: a randomised, blinded, sham-controlled trial. *Osteoarthritis Cartilage* 2018;26:1283–90.
4 Glyn-Jones S, Palmer AJR, Agricola R, *et al*. Osteoarthritis. *Lancet* 2015;386:376–87.
5 McAlindon TE, Bannuru RR, Sullivan MC, *et al*. OARSI guidelines for the non-surgical management of knee osteoarthritis. *Osteoarthritis Cartilage* 2014;22:363–88.
6 Kloppenburg M, Kroon FP, Blanco FJ, *et al*. Update of the EULAR recommendations for the management of hand osteoarthritis. *Ann Rheum Dis* 2018;2019:16–24.
7 Minten MJM, Mahler E, den Broeder AA, *et al*. The efficacy and safety of low-dose radiotherapy on pain and functioning in patients with osteoarthritis: a systematic review. *Rheumatol Int* 2016;36:133–42.
8 Bannuru RR, McAlindon TE, Sullivan MC, *et al*. Effectiveness and implications of alternative placebo treatments: a systematic review and network meta-analysis of osteoarthritis trials. *Ann Intern Med* 2015;163:365–72.
9 Arenas M, Sabater S, Hernández V, *et al*. Anti-inflammatory effects of low-dose radiotherapy. Indications, dose, and radiobiological mechanisms involved. *Strahlenther und Onkol* 2012;188:975–81.

10 Rühle PF, Wunderlich R, Deloch L, *et al*. Modulation of the peripheral immune system after low-dose radon spa therapy: detailed longitudinal immune monitoring of patients within the RAD-ON01 study. *Autoimmunity* 2017;50:133–40.
11 McKeown SR, Hatfield P, Prestwich RJD, *et al*. Radiotherapy for benign disease; assessing the risk of radiation-induced cancer following exposure to intermediate dose radiation. *BJR* 2015;88.
12 Levinson W, Kallewaard M, Bhatia RS, *et al*. 'Choosing Wisely': a growing international campaign. *BMJ Qual Saf* 2015;24:167–74.
13 Anderson GL, Judd HL, Kaunitz AM, *et al*. Effects of estrogen plus progestin on gynecologic cancers and associated diagnostic procedures: the women's health Initiative randomized trial. *JAMA* 2003;290:1739–48.
14 Roberts I, Yates D, Sandercock P, *et al*. Effect of intravenous corticosteroids on death within 14 days in 10008 adults with clinically significant head injury (MRC crash trial): randomised placebo-controlled trial. *Lancet* 2004;364:1321–8.
15 Peul WC, van Houwelingen HC, van den Hout WB, *et al*. Surgery versus prolonged conservative treatment for sciatica. *N Engl J Med* 2007;356:2245–56.